

Bag-of-Foods: Analysis of Personal Foodlogging Data

YUJI GODA, SOSUKE AMANO,
YOKO YAMAKATA, KIYOHARU AIZAWA

THE UNIVERSITY OF TOKYO

FoodLog

K. Aizawa, M.Ogawa, K. Waki, H.Takimoto, et al.
Journal of Diabetes Science and Technology 2014

- FoodLog App
 - Launched July 2013
 - Food detection & recognition

- FoodLog Data

- 2013.7.13-2017.2.6
- # of users: 114,945
- # of food records: 4,596,428
- # of photos: 2,175,328
- Nutrition information available for some foods



User ID	5901
Name	Hot black coffee
Date	2013/7/17 11:14:13
Image	
...	...

Purpose

Vectorize users' diet preference using foodlogging data

Which enables...

- Data-driven diet preference analysis
- Visualization of general bias & temporal change of diet preferences in a group / individuals

Hypothesis

Characteristics of diet preference

- Overeat
- Vegetarian
- Too much sodium
- etc.

Nutrition value represents diet preference?

Pilot Study – how average nutrition value show diet preference

User clustering based on average nutrition intake per meal

- Meal: set of foods users eat in a single eating time



- Nutrition information: 31 dimension
 - Calories
 - Fats, proteins, carbon hydroxides
 - Vitamins, minerals

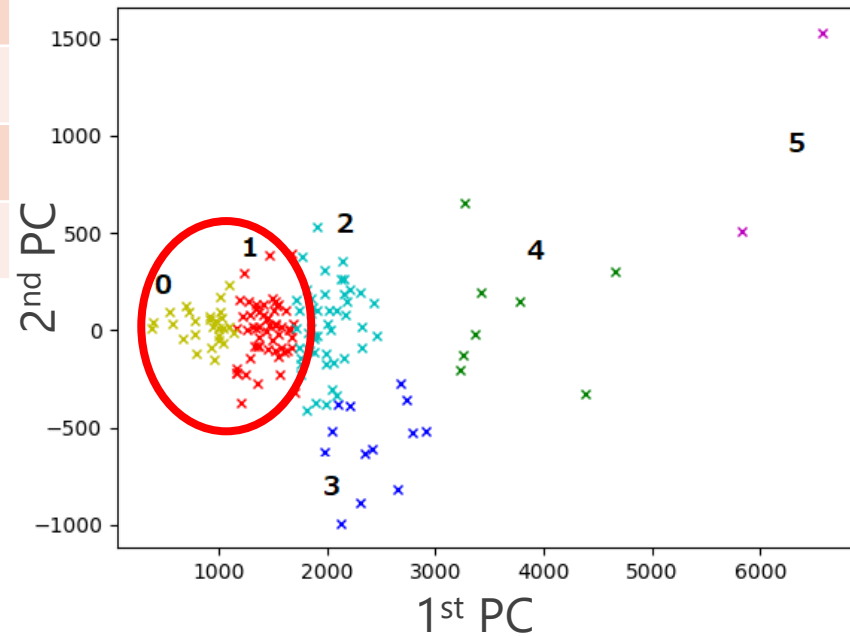
Clustering Users by Nutrition

Spare eat clusters

Typical example of meals in each cluster

Cluster	Meal
0	Jam & bread, Coffee
	Rice cake
1	Bread, Fried egg, Tomato
	Chinese soup, Rice, Natto (soy beans)

PCA & clustering result



Each point represents a user

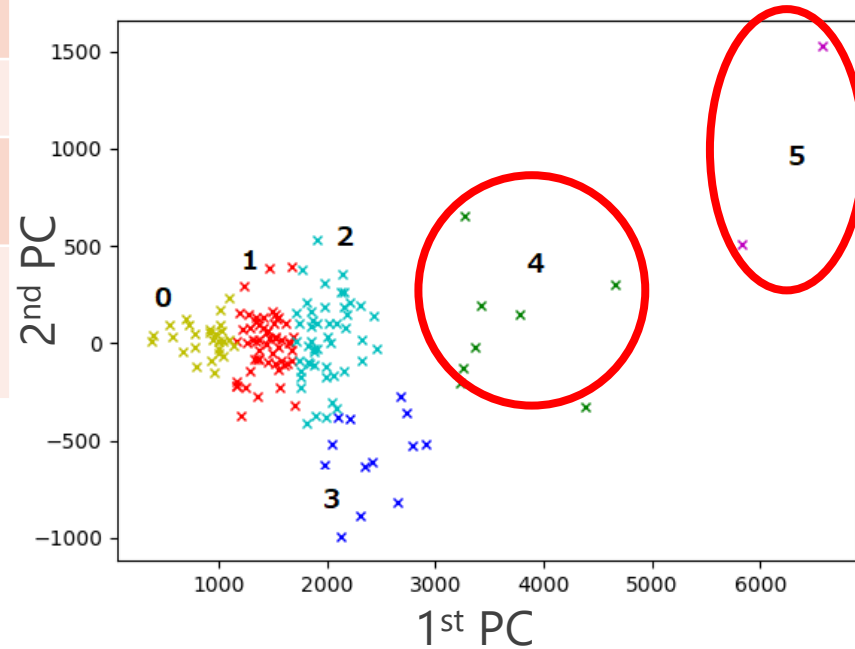
Clustering Users by Nutrition

Overeat clusters

Typical example of meals in each cluster

Cluster	Meal
4	Curry & Rice, Fresh salad, Milk
	Fried vegetables, Rice, Grilled fish, Miso soup
5	Rice, Miso soup, Fried chicken, Stewed kelp, Grilled fish, Cut cabbage
	Fried shrimp, Cut cabbage, Boiled egg, Rice, Doughnut, Rice vermicelli, Miso soup, Boiled spinach, Chili shrimp

PCA & clustering result



Each point represents a user

Conclusion of Pilot Study

Managed to discover user's diet preference characteristics such as

- Spare eater
- Overeater
- Vegetarian

just by nutrition

But, balance between each meals should be taken into account

Bag-of-Foods

Vectorize diet preferences using nutrition of each meals

- Bag-of-Foods

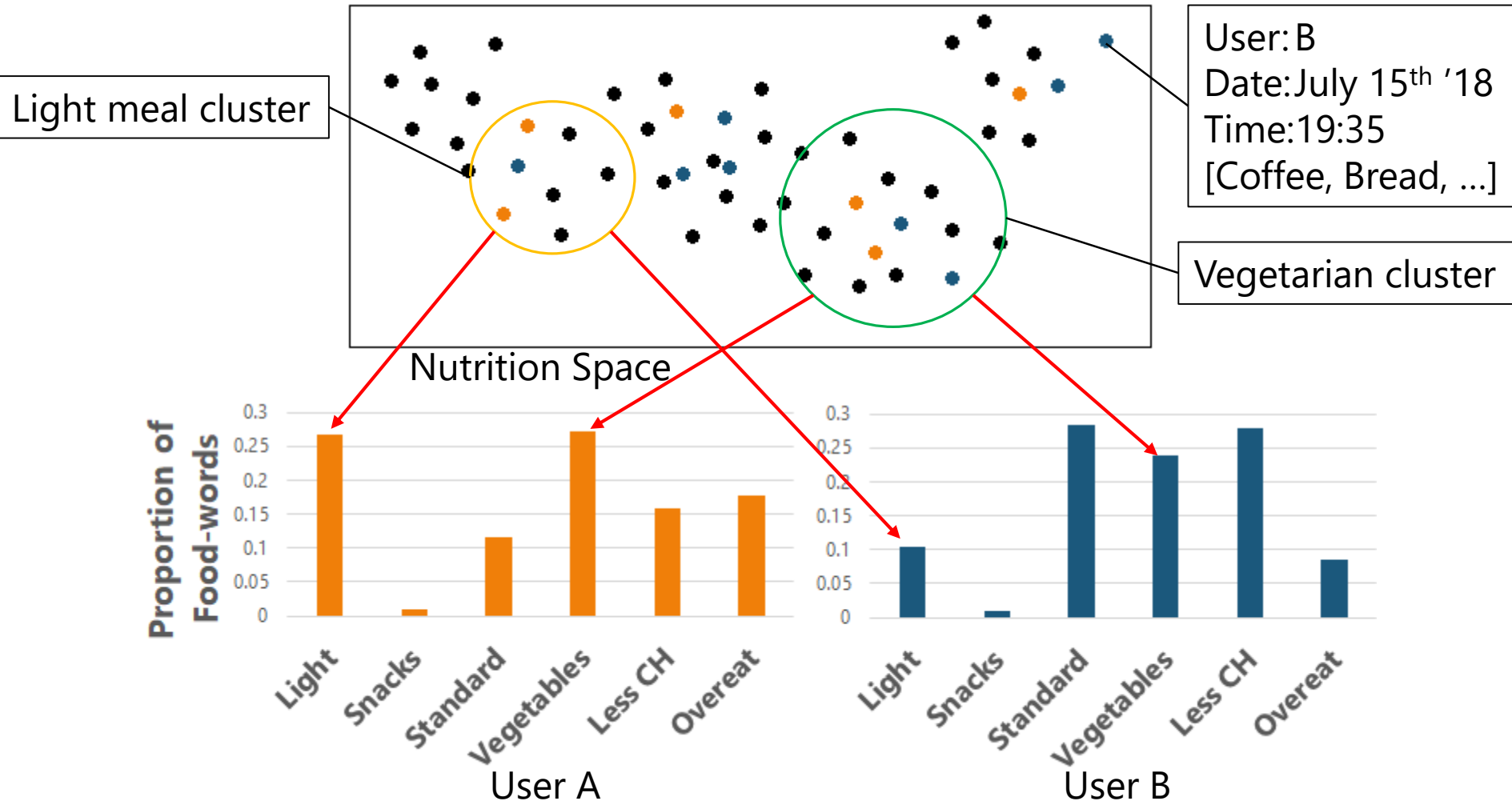
- 1-1. Cluster nutrition value of all meals by K-means

- 1-2. Treat each meal's cluster as "food-word"

- 2. Represent frequency of each food-word as diet preference

Image of Bag-of-Foods Procedure

1. Cluster all meals by nutrition value with K-means



2. Represent frequency of each food-word as diet preference

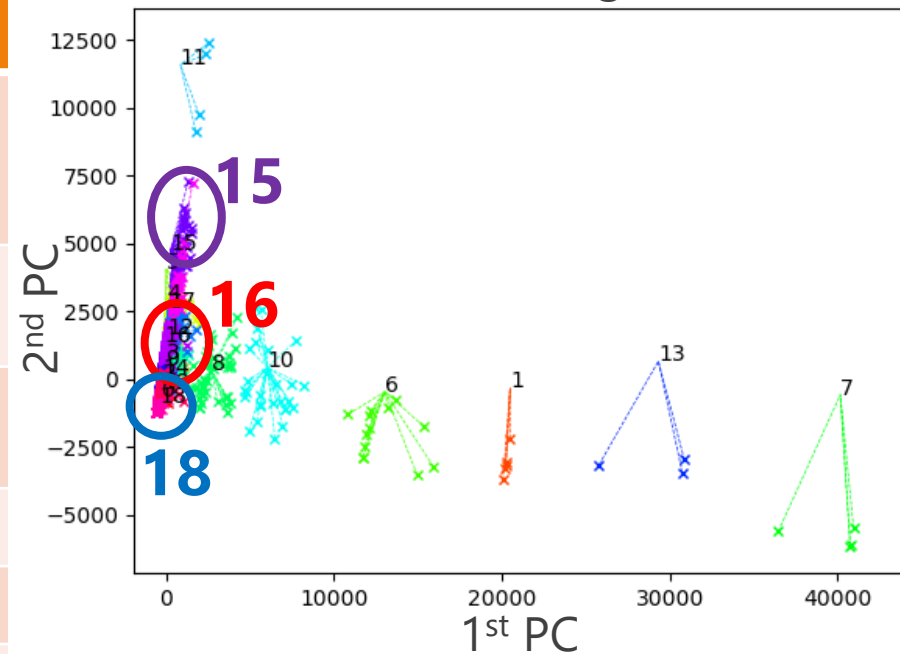
Meal Clustering Result

Clustered by amount

Typical example of meals in each cluster

Cluster	Meal
15	Curry & Rice, Boiled pork, Eggplant with miso, Fried chicken, Octopus sashimi
	Boiled spinach, Vegetable stew, Tempura, Coffee, Rice
16	Bread, Fried egg, Boiled spinach, Yogurt
	Vegetable juice, Fried fish, Tempura
18	Cheese bread
	Ice Cream

PCA & clustering result



Each point represents a meal

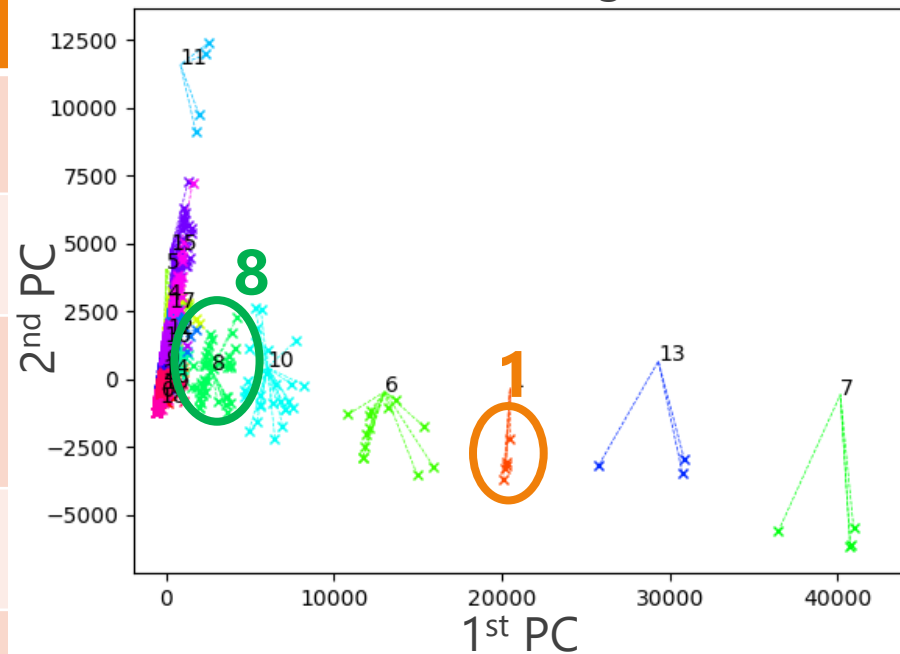
Meal Clustering Result

Clustered by nutrition balance

Typical example of meals in each cluster

Cluster	Meal
1	Boiled fish, Sake, Beer, Bread, Sashimi, Cod-chili stew, Fried tofu
	Fried turnip and chicken, Rice, Fried tofu, Yose-nabe(Stew)
	Turtle stew, Okonomiyaki, Sake, Whiskey, Fried noodle with sauce, Green salad
8	Grape, Rice, Fresh salad, Soy milk, Apple, Sausage
	Vegetable juice, Grilled fish, Miso soup, Hijiki stew, Macaroni salad, Milk

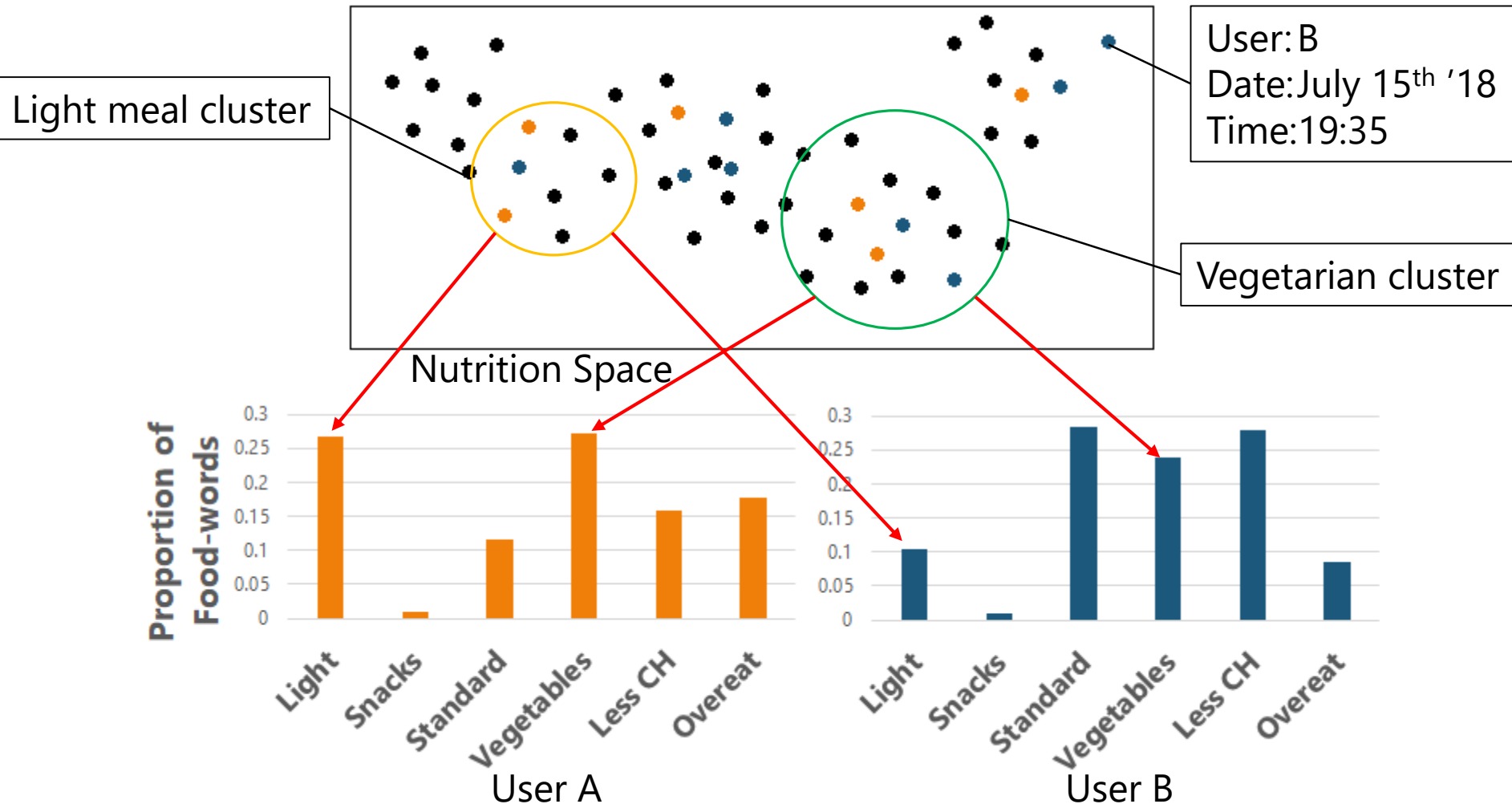
PCA & clustering result



Each point represents a meal

Image of Bag-of-Foods Procedure

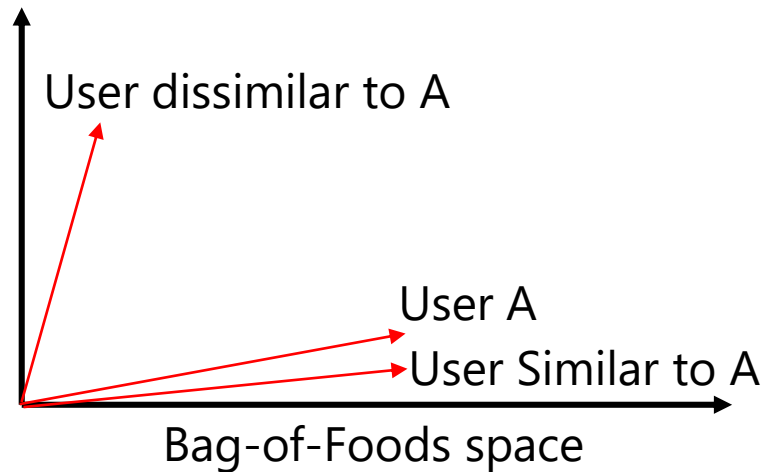
1. Cluster nutrition value of all meals by K-means



2. Represent frequency of each food-word as diet preference

Behavior of Bag-of-Foods

Chosen two users of similar BoF feature
& another user of dissimilar BoF feature



Compared their actual foodlogging data

Behavior of Bag-of-Foods

	User A	Similar User	Dissimilar User
Meals in sequence	Banana	Dry mango	Egg Sandwich
	Green salad, Rice boll	Cereals, Banana, Yogurt	Ohagi
	Nyumen noodles, Sweet beans, White raddish stew	Tomato spaghetti, Fried chicken	Baked cheese cake, Custard pudding, Black coffee
	Fried vegetables, Tuna sashimi	Fried vegetables, Tuna salad, Miso soup	Pancake, Black coffee
	Banana, Grape	Yogurt, Banana, Cereals	Punpkin stew, Miso soup, Grilled fish, Rice

Users of similar/dissimilar BoF feature and their actual foodlogging data

BoF has ability to distinguish different characteristics of diet preference

Conclusion and Future Work

- Conclusion
 - Nutrition-based clustering is effective in diet preference analysis
 - Bag-of-Foods feature managed to represent similar / dissimilar foodlogging data
- Future work
 - Utilize methods enable to handle time sequence
 - Quantitative evaluation method required