INTELLIGENT RECIPE SUGGESTION SYSTEM FOR ZERO-WASTE

TEAM MEMBERS:
Chun-hao (Larry) Chen
Deniz Genlik
Sanjay Kumar
Sevim Polat Genlik

HOW MUCH OF THE FOOD IN YOUR FRIDGE WILL YOU TOSS BEFORE IT GETS TO THE TABLE ?

Countries worldwide waste vast quantities of food annually, and the United States is one of the worst offenders.
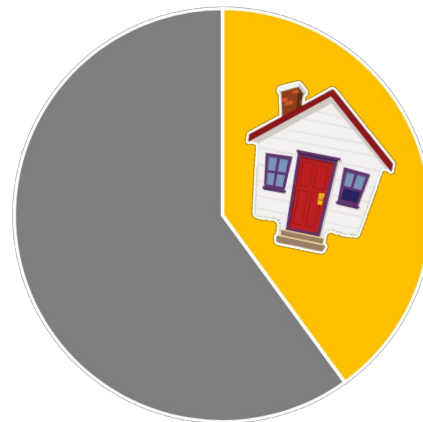
40%

of US food waste comes from individual households

Food waste will continue to deplete our economic, humanitarian, and environmental resources.

© Fotolia | Great Divide Photo

BIG IDEA : Reducing Food Waste with Smart Recipe Suggestions

Available Ingredients

RECIPES

Preferred Cuisines

Expiration Dates

USE BY: 5 JUN

BEST BEFORE: 11 MAY

MILK

User Input: Users input the ingredients available in their fridge, along with expiration dates and preferred cuisines.

Data Processing: The system processes this data to identify ingredients that need to be used soon.

Recipe Matching: The system matches these ingredients with suitable recipes from a diverse database.

Recipe Suggestions: Users receive a list of recipes that use their available ingredients and match their preferred cuisine, minimizing waste and additional purchases.

Data Cleaning

What's Cooking? Data

```python
with open('train.json') as file:
    data = json.load(file)

df = pd.DataFrame(data)
df.head()
```

	id	cuisine	ingredients
0	10259	greek	romaine lettuce, black olives, grape tomatoes...
1	25693	southern_us	[plain flour, ground pepper, salt, tomatoes, g...
2	20130	filipino	[eggs, pepper, salt, mayonaise, cooking oil, g...
3	22213	indian	[water, vegetable oil, wheat, salt]
4	13162	indian	[black pepper, shallots, cornflour, cayenne pe...

Food.com Data

```python
RAW_recipes=pd.read_csv('../data/RAW_recipes.csv')
RAW_recipes.head()
```

	name	id	minutes	contributor_id	submitted	tags	nutrition	n_steps	steps	description	ingredients
0	arriba baked winter squash mexican style	137739	55	47892	2005-09-16	['60-minutes-or-less', 'time-to-make', 'course...	51.5, 0.0, 13.0, 0.0, 2.0, 0.0, 4.0]	11	['make a choice and proceed with recipe', 'dep...	autumn is my favorite time of year to cook! th...	['winter squash', 'mexican seasoning', 'mixed ...
1	a bit different breakfast pizza	31490	30	26278	2002-06-17	['30-minutes-or-less', 'time-to-make', 'course...	[173.4, 18.0, 0.0, 17.0, 22.0, 35.0, 1.0]	9	['preheat oven to 425 degrees f', 'press dough...	this recipe calls for the crust to be prebaked...	['prepared pizza crust', 'sausage patty', 'egg...
2	all in the kitchen chili	112140	130	196586	2005-02-25	['time-to-make', 'course', 'preparation', 'mai...	[269.8, 22.0, 32.0, 48.0, 39.0, 27.0, 5.0]	6	['brown ground beef in large pot', 'add choppe...	this modified version of 'mom's' chili was a h...	['ground beef', 'yellow onions', 'diced tomato...
3	alouette potatoes	59389	45	68585	2003-04-14	['60-minutes-or-less', 'time-to-make', 'course...	[368.1, 17.0, 10.0, 2.0, 14.0, 8.0, 20.0]	11	['place potatoes in a large pot of lightly sal...	this is a super easy, great tasting, make ahea...	['spreadable cheese with garlic and herbs', 'n...
4	amish tomato ketchup for canning	44061	190	41706	2002-10-25	['weeknight', 'time-to-make', 'course', 'main-...	[352.9, 1.0, 337.0, 23.0, 3.0, 0.0, 28.0]	5	['mix all ingredients& boil for 2 1 / 2 hours ...	my dh's amish mother raised him on this recipe...	['tomato juice', 'apple cider vinegar', 'sugar...

Data Cleaning


```
['60-minutes-or-less',
 'time-to-make',
 'main-ingredient',
 'cuisine',
 'preparation',
 'for-1-or-2',
 'poultry',
 'asian',
 'chinese',
 'chicken',
 'meat',
 'number-of-servings']
```

```
['weeknight',
 'time-to-make',
 'course',
 'main-ingredient',
 'preparation',
 'main-dish',
 'pork',
 'crock-pot-slow-cooker',
 'dietary',
 'meat',
 'pork-chops',
 'equipment']
```

```
['time-to-make',
 'course',
 'preparation',
 'main-dish',
 'chili',
 'crock-pot-slow-cooker',
 'dietary',
 'equipment',
 '4-hours-or-less']
```

```
['15-minutes-or-less',
 'time-to-make',
 'course',
 'preparation',
 'low-protein',
 'healthy',
 '5-ingredients-or-less',
 'condiments-etc',
 'easy',
 'dietary',
 'low-sodium',
 'low-in-something',
 '3-steps-or-less']
```

Data Cleaning

- We extracted around 50,000 recipes with cuisine tags from 58 different cuisines.

  ⟹ ● Training Data

- Around 180,000 recipes did not have a cuisine tag.

  ⟹ ● Tags to Predict

Determining the Model for the Cuisine Predictor

Key Component: Creating a model to predict cuisines based on ingredients.
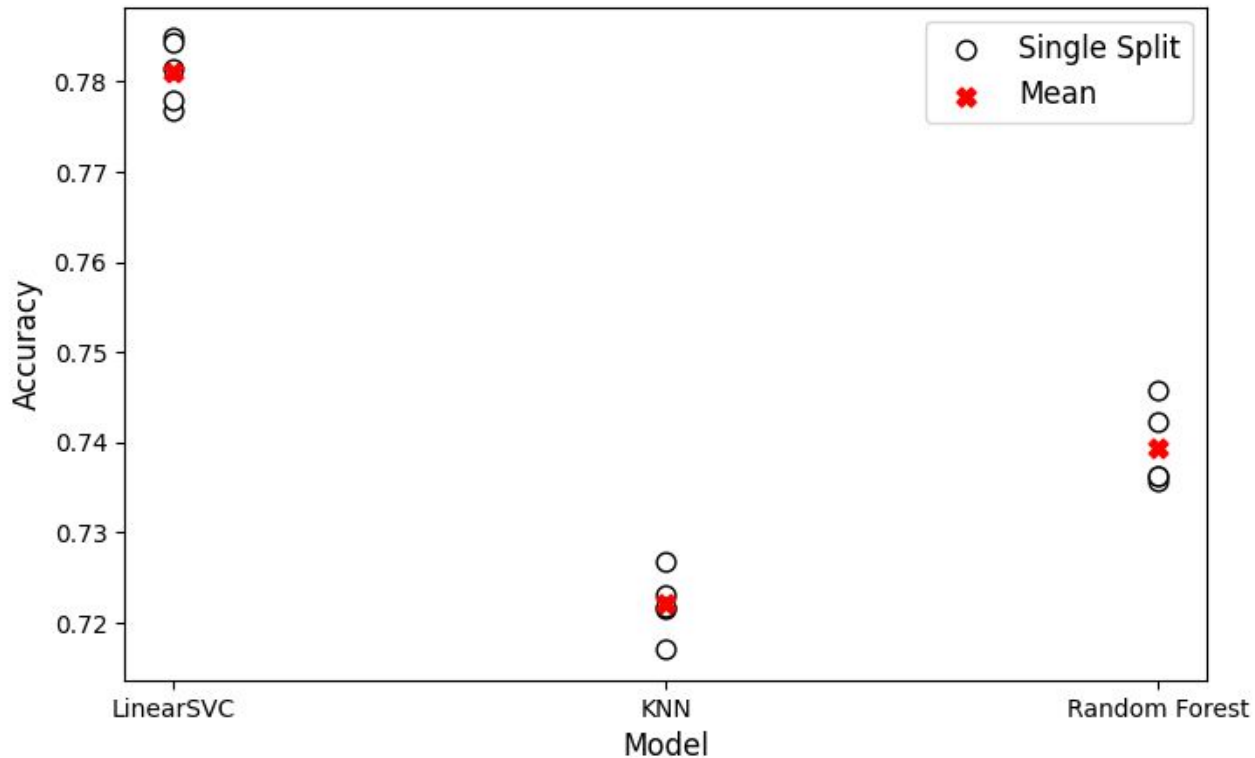
Steps:
1: Converted our data into vectors.

2: Applied cross-validation across several models on the transformed data.

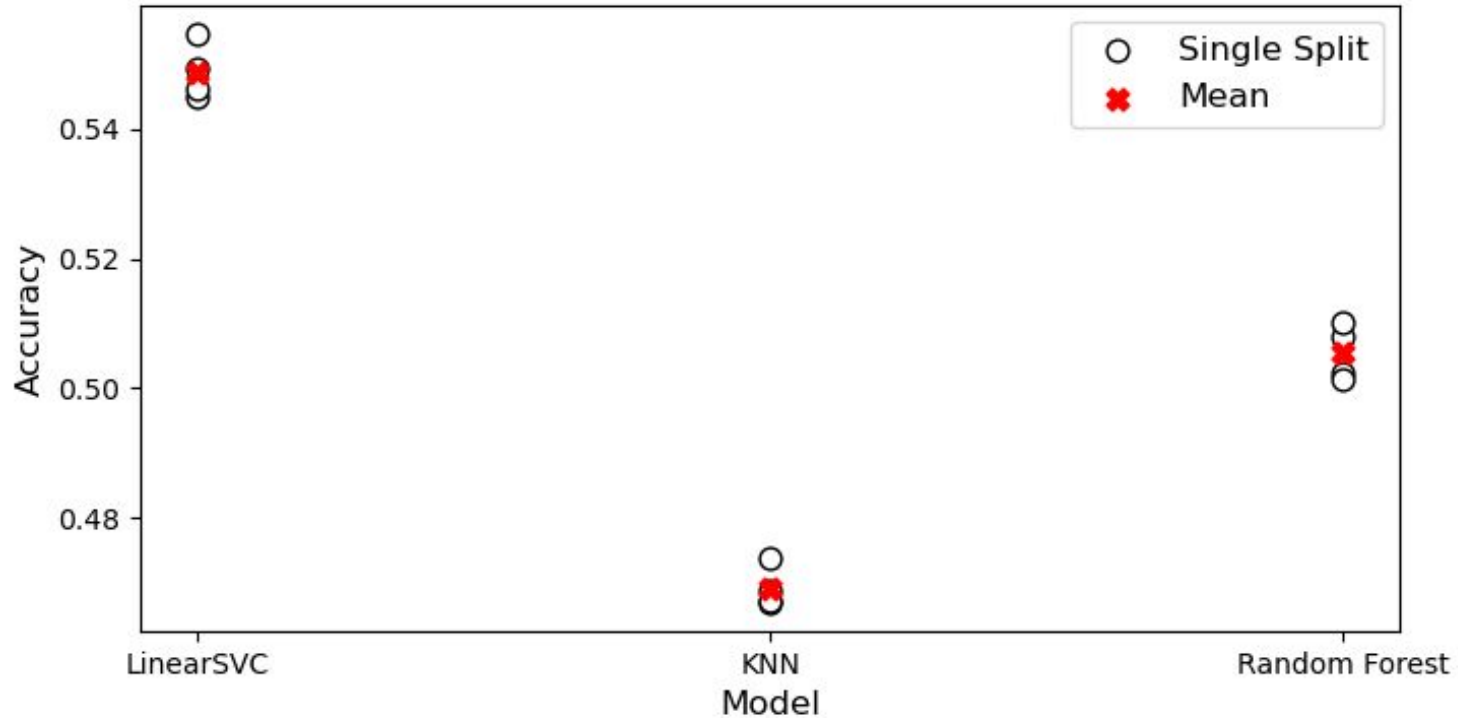**3:** Applied metrics to determine the optimal model.

4: Trained our optimal model on the dataset.

Accuracy (What's Cooking? Data)

Kaggle data with 20 cuisines

Cleaned data with 58 cuisines

Cuisine Predictor

Note: The lower performance on our cleaned data will be discussed in more detail later.

Selected Model: We train our LinearSVC model using the cleaned data.

Function: With our model, we construct the cuisine predictor function.

Example: (Cuisine Predictor)
**Input:** Cumin, Wheat, Salt, Chicken, Vegetable Oil

Output: Indian

Correlation of Cuisines

- We thought that the lower performance on food.com data is due to correlation between different cuisines.

- On the training data we calculated the correlation between each cuisine with respect to the frequency of the ingredients they use.

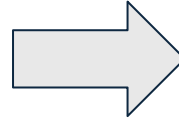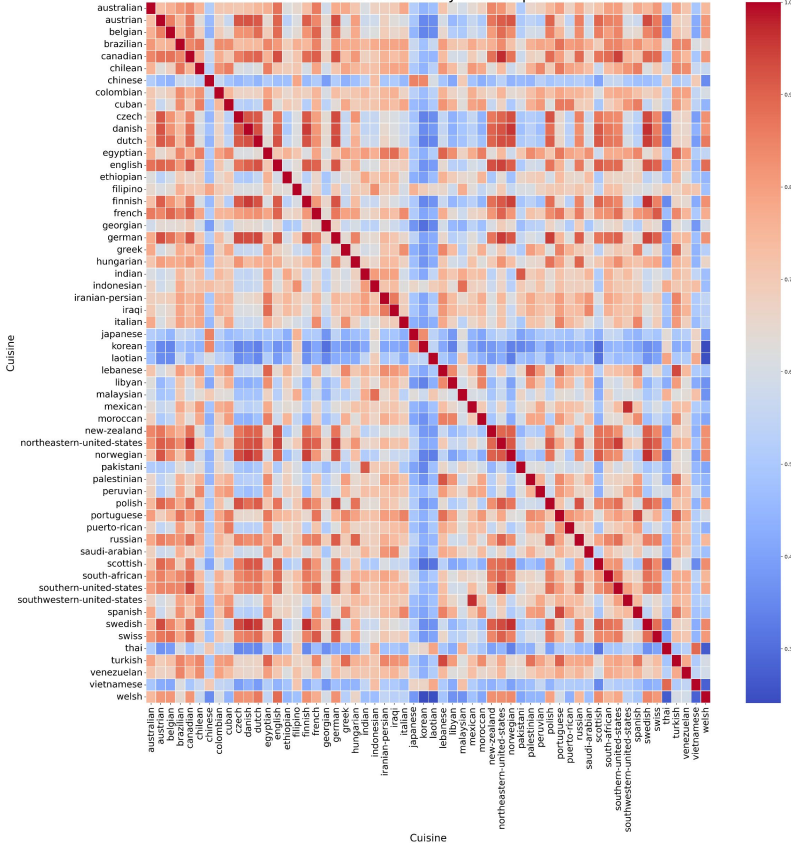| Cuisine	white cheese	roasted tomato salsa	light chunk tuna in water	mozzarella cheddar blend cheese	black tea leaf	shiso leaf	low-sugar apricot preserve	caper	pink grapefruit	tabouli mix	...	golden seedless raisin	beef blade roast	orange, juice of
australian	0	0	0	0	0	0	0	32	3	0	...	1	0	3
austrian	0	0	0	0	0	0	0	1	0	0	...	0	0	0
belgian	0	0	0	0	0	0	0	0	0	0	...	0	0	0
brazilian	0	0	0	0	0	0	0	1	0	0	...	0	0	1
canadian	0	0	0	1	0	0	1	17	4	0	...	1	2	4
chilean	0	0	0	0	0	0	0	0	0	0	...	0	0	0

Correlation of Cuisines

- We thought that the lower performance on food.com data is due to correlation between different cuisines.

- On the training data we calculated the correlation between each cuisine with respect to the frequency of the ingredients they use.
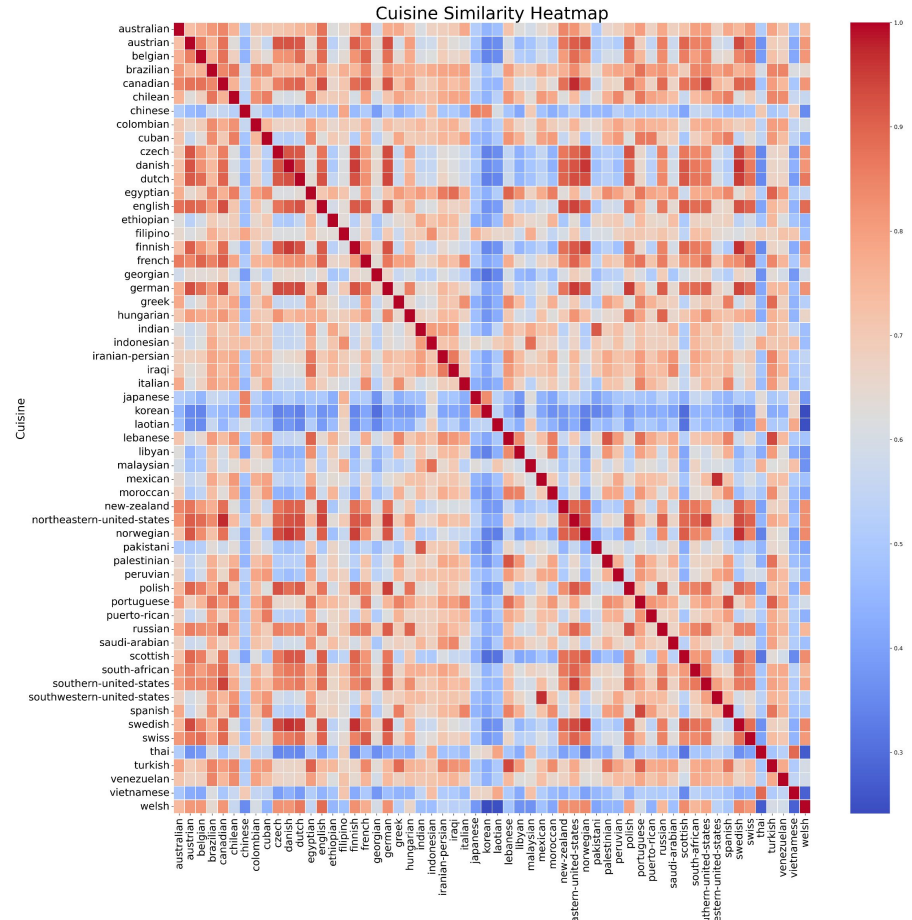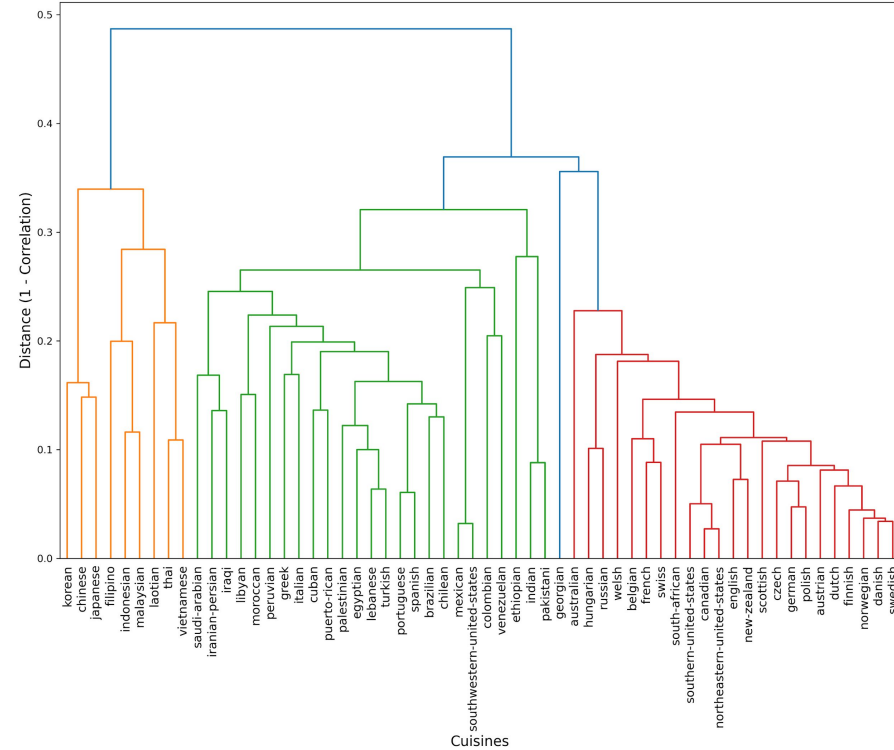

Cuisine Similarity Heatmap

Clustering of Cuisines
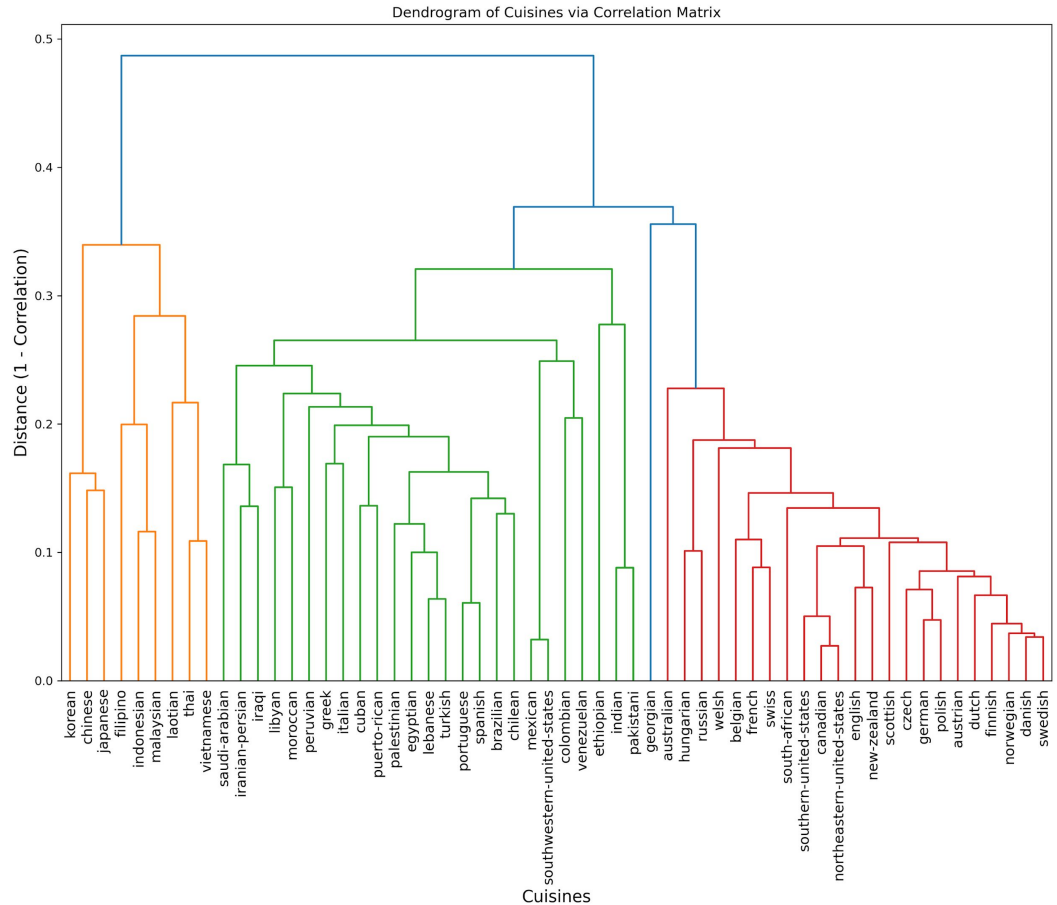
Clustering of Cuisines

Using the dendogram, we created 5 different clusters of 58 cuisines.

- 30 Clusters
- 20 Clusters
- 15 Clusters
- 10 Clusters
- 5 Clusters

Dendrogram of Cuisines via Correlation Matrix

Clustering of Cuisines

```
{'1': ['chinese', 'japanese', 'korean'],
 '2': ['filipino', 'indonesian', 'malaysian'],
 '3': ['laotian', 'thai', 'vietnamese'],
 '4': ['brazilian', 'chilean', 'cuban', 'egyptian', 'greek', 'iranian-persian', 'iraqi', 'italian',
       'lebanese', 'libyan', 'moroccan', 'palestinian', 'peruvian', 'portuguese', 'puerto-rican',
       'saudi-arabian', 'spanish', 'turkish'],
 '5': ['mexican', 'southwestern-united-states'],
 '6': ['colombian', 'venezuelan'],
 '7': ['indian', 'pakistani'],
 '8': ['ethiopian'],
 '9': ['australian', 'austrian', 'belgian', 'canadian', 'czech', 'danish', 'dutch', 'english',
       'finnish', 'french', 'german', 'hungarian', 'new-zealand', 'northeastern-united-states',
       'norwegian', 'polish', 'russian', 'scottish', 'south-african', 'southern-united-states',
       'swedish', 'swiss', 'welsh'],
 '10': ['georgian']}
```
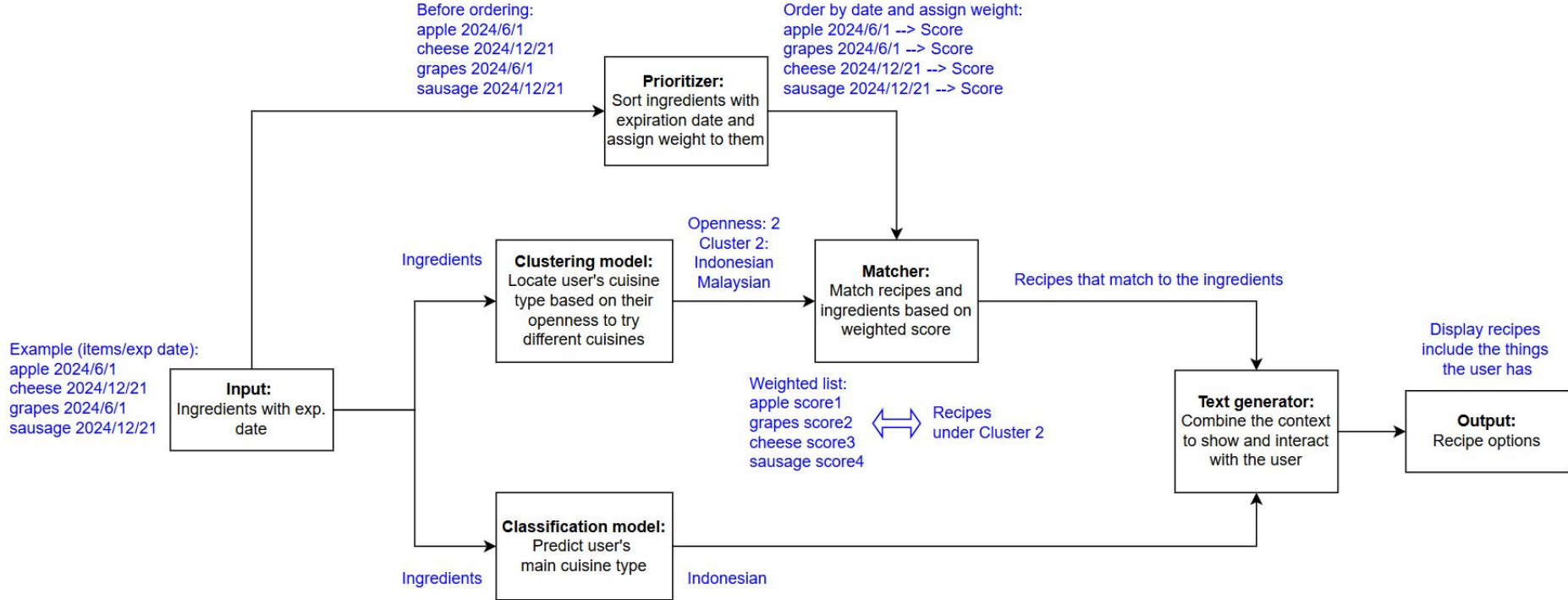
$$\underbrace{\begin{bmatrix} R_{11} & \dots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{m1} & \dots & R_{mn} \end{bmatrix}}_{\text{Binary Recipe Matrix}} \cdot \underbrace{\begin{bmatrix} I_1 \\ \vdots \\ I_n \end{bmatrix}}_{\substack{\text{Weighted} \\ \text{Ingredient} \\ \text{Vector}}} = \underbrace{\begin{bmatrix} S_1 \\ \vdots \\ S_m \end{bmatrix}}_{\substack{\text{Scores of} \\ \text{Recipes}}}$$

$$I_j = \frac{1}{(0.01 + \text{shelf life})}$$

Flow Chart

Intelligent Recipe Suggestion System For Zero-Waste

Before ordering:
apple 2024/6/1
cheese 2024/12/21
grapes 2024/6/1
sausage 2024/12/21

Order by date and assign weight:
apple 2024/6/1 --> Score
grapes 2024/6/1 --> Score
cheese 2024/12/21 --> Score
sausage 2024/12/21 --> Score

Prioritizer:
Sort ingredients with expiration date and assign weight to them

Ingredients

Openness: 2
Cluster 2:
Indonesian
Malaysian

Clustering model:
Locate user's cuisine type based on their openness to try different cuisines

Matcher:
Match recipes and ingredients based on weighted score

Recipes that match to the ingredients

Example (items/exp date):
apple 2024/6/1
cheese 2024/12/21
grapes 2024/6/1
sausage 2024/12/21

Input:
Ingredients with exp. date

Weighted list:
apple score1
grapes score2
cheese score3
sausage score4

Recipes
under Cluster 2

Display recipes include the things the user has

Text generator:
Combine the context to show and interact with the user

Output:
Recipe options

Ingredients

Classification model:
Predict user's main cuisine type

Indonesian

Demo

Summary

Goal: Construct an intelligent recipe suggestion system for zero-waste which considers users' preferences.

Final Results: Through preprocessing, clustering, and creating a metric for weighting approaching expiration dates, we achieved our goals.

Future Directions

For improved user experience, we plan on implementing the following features to our program:

- Weighting recipes based on users likes and dislikes

- Adding an estimation of savings earned by utilizing foods with an approaching expiration date

- Provides recipes which minimizes additional costs of ingredients

- Algorithm to maximize total shelf life of ingredients