PROJECT REPORT

FOOD RECOMMENDATION SYSTEM

Team Members: Kondrakunta Ramya, Marathi Vinitha Joyce

INTRODUCTION

Food has always been of great importance in various cultures and plays a significant role in festivals and celebration in many religions. Food was, is and will continue to help build connections between people. The motivation behind this project is to provide food recommendations with ease. Inverse Cooking: Recipe Generation From Food Images[4] inspired this project to recommend items along with their ingredients taking either an image or set of ingredients from the user.

Contributions

- Documentation: Kondrakunta Ramya, Marathi Vinitha Joyce
- Data preprocessing: Kondrakunta Ramya
- Recommendations using images: Marathi Vinitha Joyce
- Recommendations using text: Kondrakunta Ramya

RELATED WORK

Recommendation systems are mainly of two types: collaborative and content-based. The former is dependent on user reviews and the latter is based on similarity of the content. Coming to food recommendation systems, multiple work in building these systems have been done.[1] The main goal for a lot of these systems has been to recommend food considering nutritional value, calorie content as well as the amount of consumption. Various approaches to tackle each of these issues have been discussed in A survey on AI nutrition recommender systems.[2]

Building these systems is challenging taking into consideration all the various parameters. Being able to recommend ingredients of the item based on a user's preference by taking either an image or set of ingredients. Similar food dishes will be recommended. Both image and text approaches are analyzed to see how they respond to various clustering and dimensionality reduction methods.

BACKGROUND INFORMATION

Clustering

Clustering is the task of dividing the population or data points into a number of groups such that data points in the same groups are more similar to other data points in the same group than those in other groups.

T-Distributed Stochastic Neighbor Embedding (t-SNE):

It's a non-linear, unsupervised technique for data exploration and visualization of high-dimensional data. t -SNE offers a sense of how data is organized in a high-dimensional space. The t-SNE technique calculates a similarity measure between pairs of instances in high and low dimensional space. It then leverages a cost function to maximize these two similarity measures.

K-means

The KMeans algorithm clusters data by attempting to split samples into n clusters of equal variance while minimizing the inertia or within-cluster sum-of-squares criterion. The number of clusters must be specified in this technique.

Principal Component Analysis (PCA)

PCA is utilized in exploratory data analysis and predictive model development. It's often used for dimensionality reduction, where each data point is projected onto only the first few principal components to produce lower-dimensional data while preserving as much variance as feasible.

Transformers

A transformer is a deep learning model that uses the self-attention mechanism to weigh the importance of each element of the input data uniquely. It is largely utilized in natural language processing and computer vision applications (CV).

VGG-16

VGG-16 is a convolutional neural network that is 16 layers deep. You can load a pre trained version of the network trained on more than a million images from the ImageNet database.

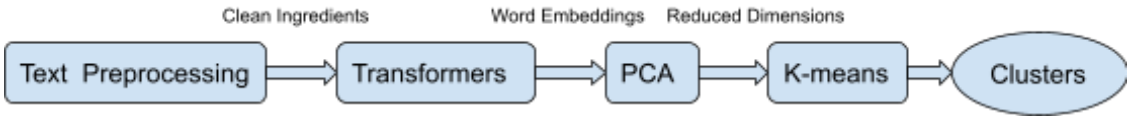
PROPOSED APPROACH

Two approaches are being explored in this project.

1. Recommendation using Ingredients (Text data)
2. Recommendations using Food Images (Image data)

Recommendation using Ingredients

This approach involves taking the ingredients of an item and recommends top 5 similar food items. The following was implemented to achieve results: NLP tasks, T-SNE for initial visualization of the data, dimensionality reduction using PCA to perform K-means clustering.

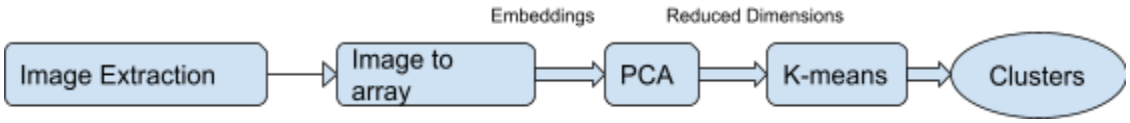


Text preprocessing is an essential step to cluster and arrive at good recommendations.
The dataset consists of nearly one million instances resulting in a huge number of ingredients. With the help of the spacy library, after the basic cleaning of the text, nouns were extracted to derive various ingredients. Three word ingredients were considered and ingredients exceeding more than three words were discarded.[3][4]
The item's ingredients were converted to word vectors using the Sentence Transformer from Hugging Face. These word embeddings were used in clustering. T-SNE was used for visualization and K-means clustering was used for recommending similar items.
For visualization purposes, customized categories were created. The items were categories into cuisines based on how similar their ingredients are to the cuisines defined. 8 cuisines were defined with their cuisine specific ingredients namely: Indian, Greek, Chinese, Italian, Japanese, Korean, Mexican, Thai/Vietnamese, Other.
With the help of PCA, word embeddings were reduced to 2 components and K-means was performed. The elbow method was used to find the optimal cluster value and silhouette coefficient was estimated for different

cluster values. The predicted clusters on the test data were mapped to the top 5 similar items, here the similarity was determined using cosine similarity.

Recommendations using Food Images

This process involves clustering the images into similar clusters based on similarity of images into 200 clusters. The tasks performed in this process are image extraction, creating embeddings by converting images to 224x224 array format, Dimensionality reduction, Clustering using K-means.



Here, we classify the images based on similarity of the images so it can classify food in categories. Some categories that were created are: pizzas, cakes, cookies, drinks, salads, etc.

EXPERIMENTS

Dataset

The dataset used is the Recipe1M, it is the largest open dataset with 887,706 images and recipes. It contained two json files: layer1 and layer2. Layer1.json contained all text related data and Layer2.json had the links to images. Both text and image .json files can be mapped together with the “id” for each data instance.

Results

Using Ingredients

Customized Cuisines: The cuisines were defined to have the cuisine specific ingredients

```
Indian = ["garam masala","turmeric","cumin seed","green chilies","tamarind","curry leaves","coriander","fenugreek","bay leaves","paneer","ghee","basmati rice"]
Greek = ["feta cheese","oregano","phyllo dough","greek yogurt","walnuts","fava beans","sardines","anchovies"]
Chinese = ["hoisin sauce","oyster sauce","sesame oil","corn starch","peanut oil","soy sauce","white rice vinegar","mushrooms","sichuan","chilli bean sauce"]
Italian = ["parmesan","basil","olive oil","black pepper","pasta","wine","cherry tomato","oregano","balsamic vineger","garlic","capers"]
Japanese = ["sake","mirin","rice vinegar","soy sauce","scallions","miso","wakame","bonito flakes","kombu","shichimi","nori","wasabi"]
Korean = ["gochujang","kimchi","sesame seed oil","sesame seeds","sesame oil","tofu","soy sauce","kochujang","ginger","rice wine","garlic","rice"]
Mexican = ["corn tortillas","salsa","flour tortillas","beans","avocado","pinto beans","cilantro","limes","cheese","tomatos","corn","poblano pepper","chipotles"]
Thai_Vietnamese= ["fish sauce","lemongrass","coconut milk","beansprouts","lime juice","peanuts","shallots","shallot oil","scallions","chiles","thai basil","shrimp paste","jasmine rice","rice vermicelli"]
```

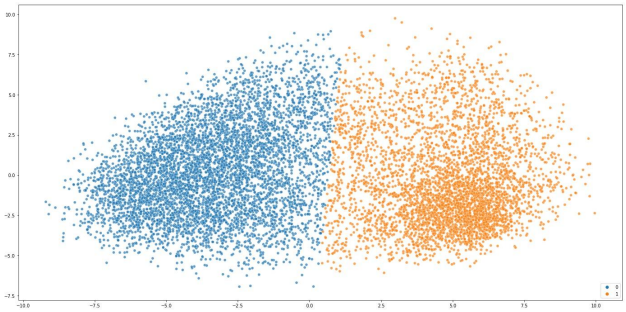
Visualizations T-SNE

Data size= 10,000
Perplexity=30,Iteration=5000

Plot 1

K-means

Clusters=2,3,4,5



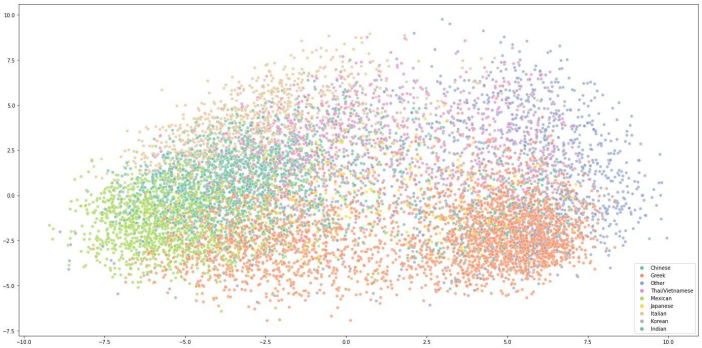
Plot3

Plot4

Plot5

Plot6

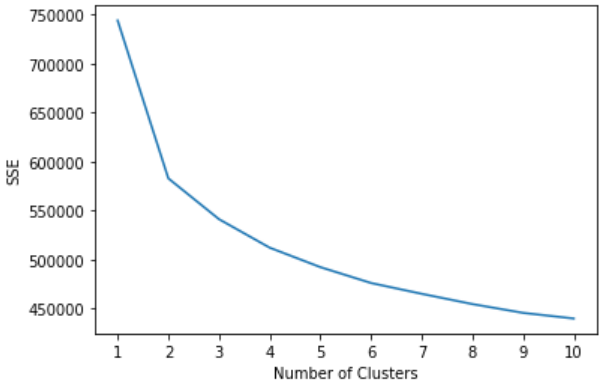
Plot after PCA (based on Cuisines)



Plot7

We can observe here that *Greek* food along with the *Other* are mainly clustered to the right side. Cuisines such as *Chinese*, *Korean* and *Indian* are mostly clustered to the left side. *Japanese* and *Thai/Vietnamese* seem to be scattered on both sides.

Elbow Plot for picking the right cluster value



Plot 8

Recommendations using K-means

When Clusters=9

Given item: ***Gluten-free Vanilla Bean Cupcakes***

Recommended items and their ingredients:

Macadamia Butter/Cranberry Cookies:*['macadamia nuts', 'butter', 'sugar', 'vanilla', 'egg', 'flour', 'soda', 'salt', 'nutmeg', 'cranberries'],*

Spicey Nut Cake (Also Gluten Free):*['butter', 'sugar', 'weight cream cheese', 'eggs', 'almond flour', 'pecans', 'guar gum', 'powder', 'soda', 'ground cinnamon', 'ground ginger', 'ground allspice', 'ground', 'vanilla extract', 'salt', 'pineapple pineapple juice', 'raisins', ''],*

Fluffy Strawberry Meringue Pie:*['egg whites', 'cream tartar', 'sugar', 'saltine crackers', 'pecans', 'vanilla extract', 'strawberries', 'marshmallows', 'topping', 'food'],*

Healthy Cranberry Walnut Muffins:*['oats', 'flour', 'powder', 'salt', 'cranberries', 'walnuts', 'yogurt', 'applesauce sugar type', 'sugar', 'eggs', 'cinnamon', 'vanilla'],*

Gluten Free Lemon Poppyseed Muffins:*['sugar', 'lemon', 'purpose flour blend', 'ground almonds', 'teff flour', 'powder', 'soda', 'salt', 'greek yogurt', 'eggs', 'vanilla paste', 'butter', 'poppyseeds', 'milk']*

When Clusters=3

Given item: ***Pita Bread***

Recommended items and their ingredients:

Oeufs Enterallies:*['eggs', 'bacon', 'butter', 'flour', 'cream', 'milk', 'cheddar cheese', 'clove', 'thyme', 'marjoram', 'basil', 'parsley', 'bread'],*

Chef Black Rayne's Curry Mac & Cheese (from box):*['mac cheese', 'cream', 'bunch parsley', 'powder', 'seasoning', 'pepper', 'yogurt'],*

Amberjack Simmered in Miso:*['pieces buri', 'egg daikon', 'water', 'sake', 'sugar', 'soy sauce', 'miso'],*

Apple and Sweet Potato Towers:*['potatoes', 'apples', 'nutmeg', 'cinnamon', 'sugar', 'butter', 'corn flour', 'biscuits', 'zest lemon', 'greek yogurt'],*

Greek Cornbread With Cheese - Bobota Me Tyri:*['greek yogurt', 'oil', 'eggs', 'lemon juice', 'cornmeal fine', 'purpose flour', 'sugar', 'soda', 'powder', 'salt', 'feta cheese', 'butter cut pieces']*

Using Images

Sample cluster: This cluster has drinks in it.

Formed Cluster below can be labeled as Desserts. Although a few items like Moroccan Lamb and Cranberry Onion Chicken were not expected to be in this cluster, the remaining items can be classified under desserts.

```
clusters[7]
```

```
['Baked Berry Oatmeal Goodness.png',
 'Cafe-Style Puffed Wheat Squares.png',
 'Snowcapped Cherry Nibbles.png',
 'Better-Than-Regular Brownies (Christmas Style!).png',
 'Moroccan Lamb with Shiraz Honey Sauce.png',
 'Cranberry Onion Chicken.png',
 'Chinese Spareribs.png',
 'Turkey Bacon Cornbread Muffins.png',
 'Tropical Kumquat Cake.png',
 'Eclair Cake.png',
 'Eggplant Torta.png',
 'Russian Stuffed Meatloaf.png',
 'No Bake " Chocolate Chip Cookie Pie " ( made with Chips Ahoy)....png',
 'Candy Cane KISS Blondies.png',
 'Fluffernutter SMores Brownie Bars.png',
 'Chocolate Nutella Molten Lava Cake.png',
 'Dark Chocolate, Cranberry and Macadamia Bars.png']
```

Formed Cluster below can be labeled as Salad. Although a few items like Strawberry Dressing and Grilled Caprese Chicken were not expected to be in this cluster, the remaining were Salads as shown.

```
[ ] clusters[13]
```

```
['Eggplant Tomato Salad.png',
 'Strawberry Dressing.png',
 'Stir-fried Tofu With Carrots and Red Peppers.png',
 'Red, White and Blue Salad.png',
 'Grilled Scallops with Parsley Salad.png',
 'Zucchini Coleslaw.png',
 'Greek Grilled Chicken Salad.png',
 'Brie, Pear, Cranberry, and Walnut Salad.png',
 'Tomato and Basil with Lemon Flavored Dressing.png',
 "Grandma's French Dressing.png",
 'Kale With Garlic N Tomatoes.png',
 'Grilled Caprese Chicken with a Balsamic Reduction.png',
 'Easy Cajun Chicken Caesar Salad.png',
 'Chicken Nicoise Salad.png',
 'Roasted Broccoli and Tomatoes.png',
 'Potato, Italian Sausage, and Arugula Salad.png']
```

CONCLUSIONS AND FUTURE WORK

Summary, Future Work and Challenges

The recommendations given an image or text are satisfactory. We have not just recommended similar items but have also given ingredients that these items might possibly have.

Scope for this project is that given an image, predict the ingredients directly from the image and give recommendations based on these predicted images. Further, recipes can be derived from these ingredients (inspired from Inverse Cooking: Recipe Generation From Food Images[4]

Working with a huge dataset is very difficult without the right computational power. It can be very expensive. There will always be a compromise on either the data size or the accuracy of the models.

REFERENCES

[1] Trang Tran, T.N., Atas, M., Felfernig, A. *et al.* An overview of recommender systems in the healthy food domain. *J Intell Inf Syst* 50, 501–526 (2018). https://doi.org/10.1007/s10844-017-0469-0

[2] Thomas Theodoridis, Vassilios Solachidis, Kosmas Dimitropoulos, Lazaros Gymnopoulos, and Petros Daras. 2019. A survey on AI nutrition recommender systems. In Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '19). Association for Computing Machinery, New York, NY, USA, 540–546. https://doi.org/10.1145/3316782.3322760

[3] A. Salvador et al., "Learning Cross-Modal Embeddings for Cooking Recipes and Food Images," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 3068-3076, doi: 10.1109/CVPR.2017.327.

[4] Salvador, Amaia & Drozdzal, Michal & Giró-i-Nieto, Xavier & Romero, Adriana. (2019). Inverse Cooking: Recipe Generation From Food Images. 10445-10454. 10.1109/CVPR.2019.01070.