

# Culinary Compass

*Flavour Meets Wellness*

**Capstone Project**

Neuefische Data Practitioner Bootcamp

07 April 2025



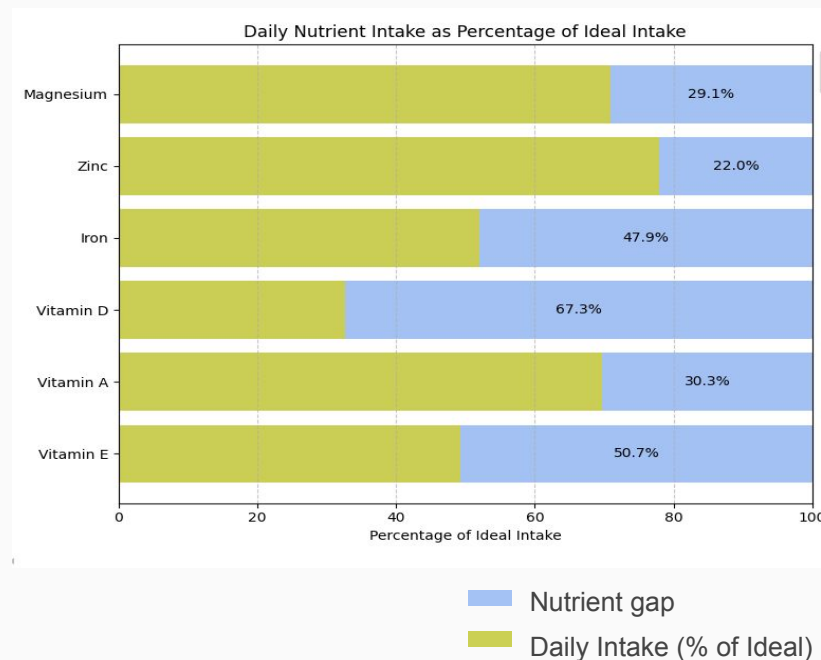
# Hidden Hunger – The Global Challenge

**Hidden Hunger:** A form of malnutrition where individuals have enough calories but lack essential vitamins and minerals. Despite food abundance, many diets lack essential micronutrients.

Affects over 2 billion people globally.

Leads to:

- Impaired cognitive development.
- Increased susceptibility to infections.
- Chronic diseases.



# Culinary Compass - Flavor Meets Wellness

## Empowering Individuals with Informed Choices

Personalized food and recipe recommendations based on nutritional needs.

## Goal

## Build a **Nutrition-Based Food and Recipe Recommendation system** that:

- Recommends food items based on nutritional deficiencies.
- Recommends recipes based on the suggested food items using **NLP**.



# Dataset Overview

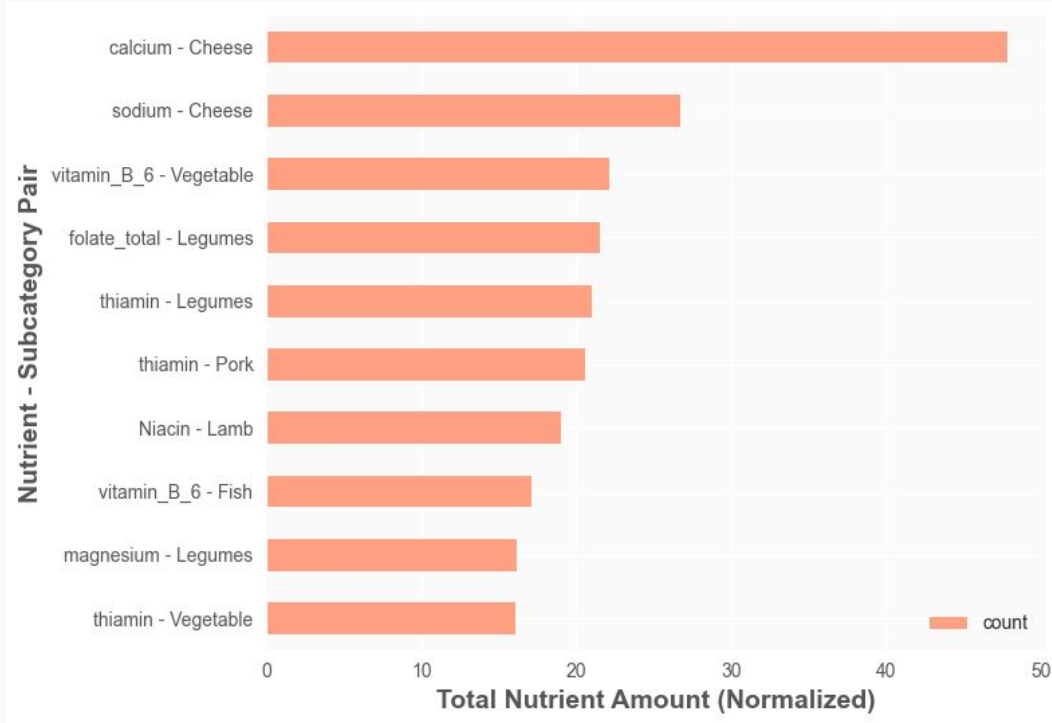
Dataset	Food	Recipe
Source	<u>USDA Website</u>	Food.com <u>Kaggle</u>
Size	1166	522K (Trained 20K)
Features	36 Features 22 Features utilised	28 Features 15 Features utilised
Focus	Recommend Food Items	Recommend Recipes



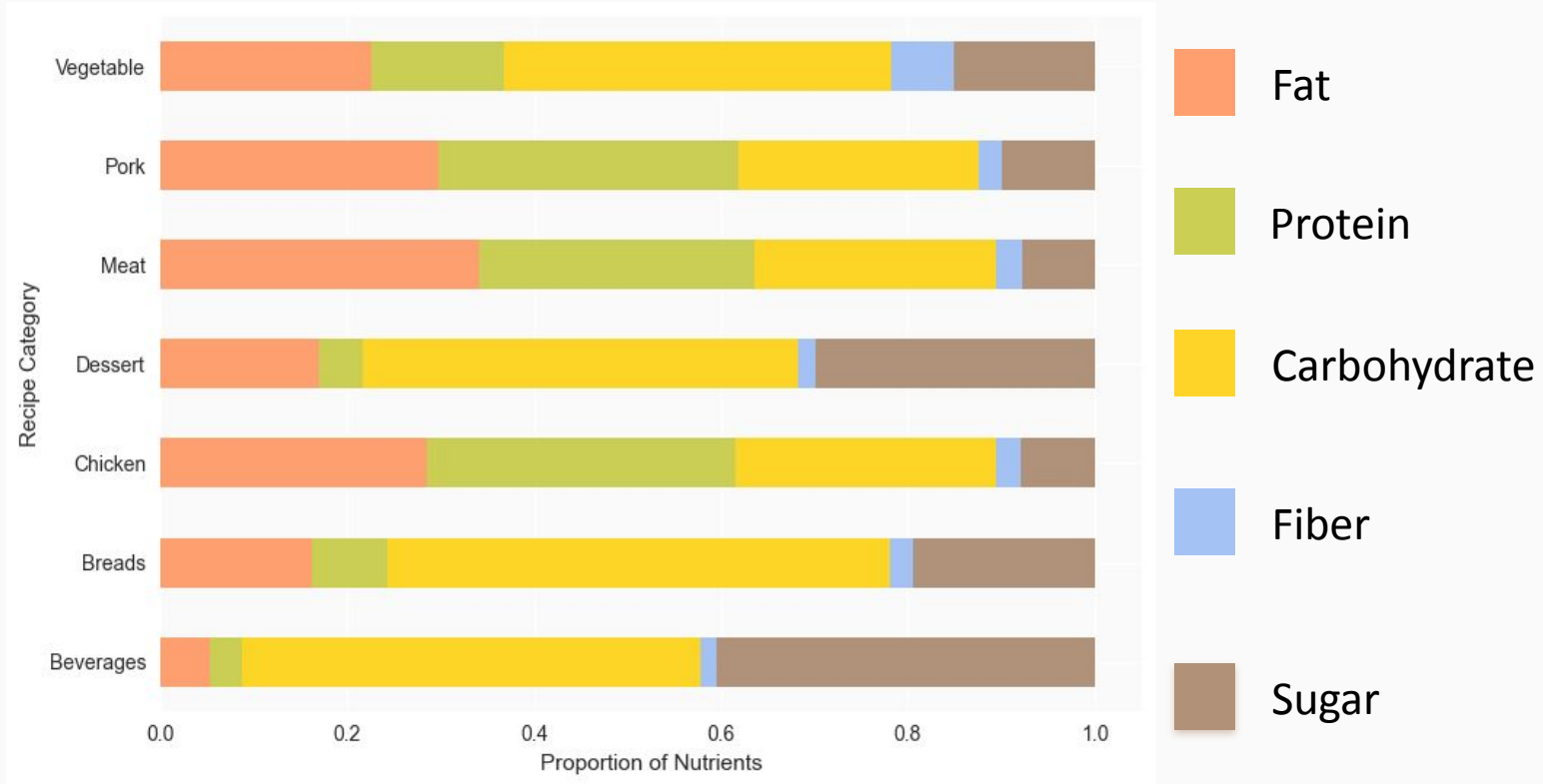
# Data Preprocessing

Dataset	Cleaning	Preprocessing
Food	<ul style="list-style-type: none"><li>● Processed and cleaned food item names</li><li>● Converted micro nutrients like iron, vitamin D, Vitamin C etc to same unit (milligrams).</li></ul>	<ul style="list-style-type: none"><li>● Scaling - MinMax Scaler</li><li>● Creating Categories and Subcategories</li></ul>
Recipe	<ul style="list-style-type: none"><li>● Removing rows that has null values for macronutrients</li><li>● Converted macronutrients like protein, fat, carbohydrate etc to same unit (grams).</li></ul>	<ul style="list-style-type: none"><li>● Batch processing of data to generate embeddings</li><li>● Embedding were generated for Recipe ingredients.</li></ul>

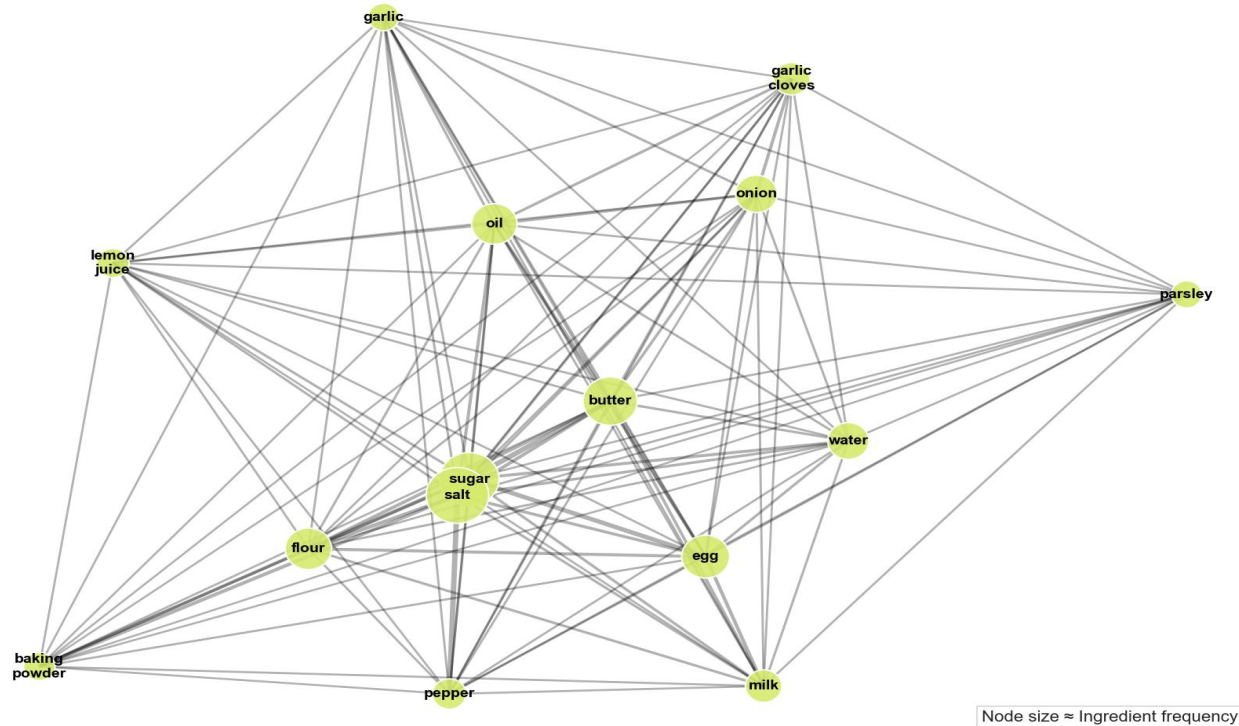
# Top 10 Nutrient Subcategory Associations



# Distribution Of Macronutrients

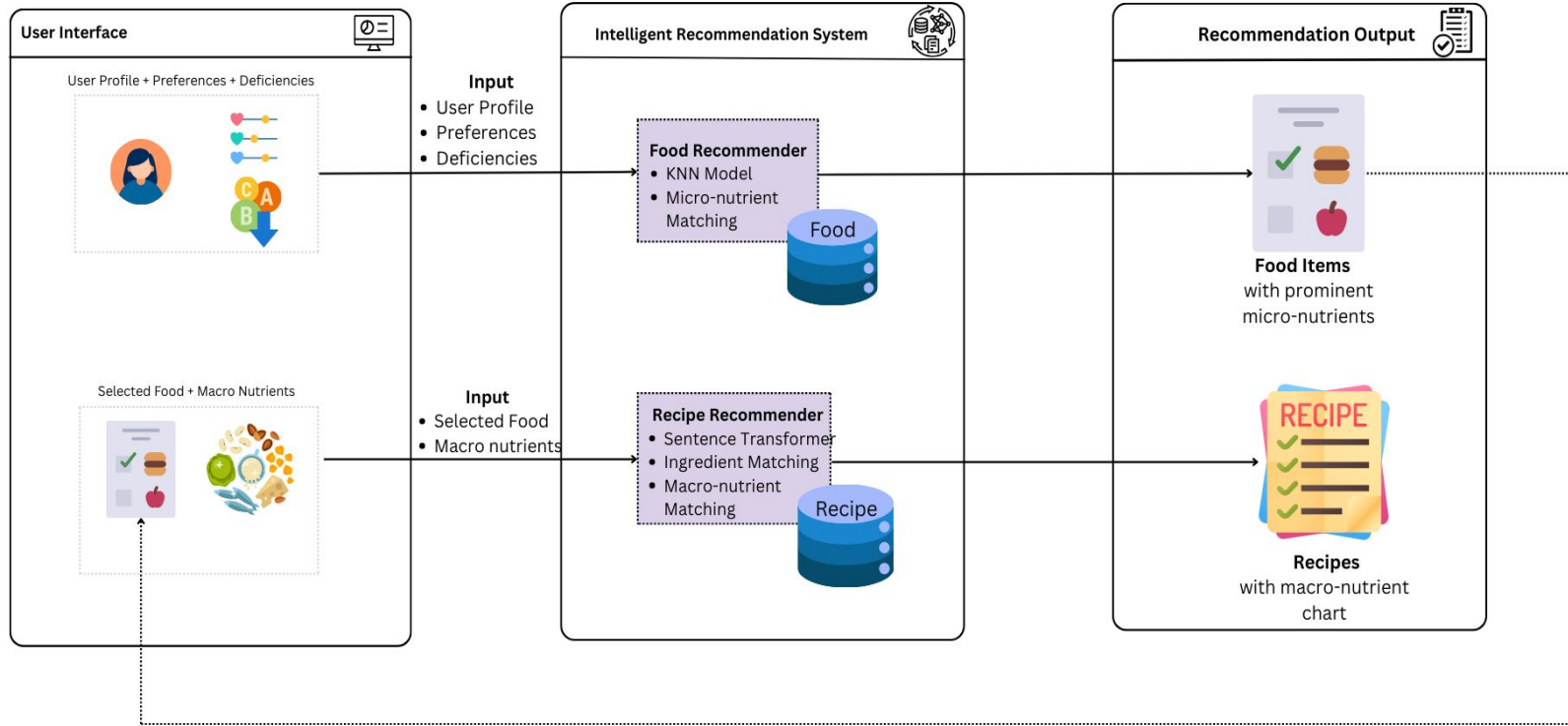


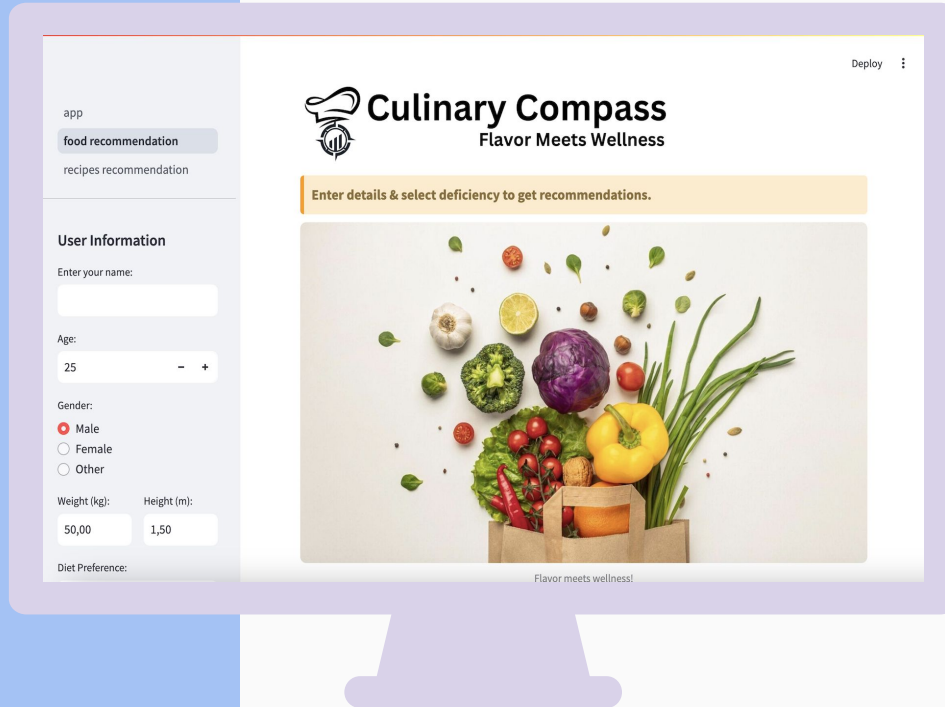
# Recipe Ingredients Co-occurrence





# System Architecture





The screenshot shows a web application titled "Culinary Compass" with the tagline "Flavor Meets Wellness". The interface is divided into a left sidebar and a main content area. The sidebar contains a navigation menu with "app", "food recommendation" (highlighted), and "recipes recommendation". Below the menu is a "User Information" section with fields for "Enter your name:", "Age:" (set to 25), "Gender:" (with radio buttons for Male, Female, and Other), "Weight (kg):" (set to 50,00), "Height (m):" (set to 1,50), and "Diet Preference:". The main content area features a header with the app's logo and tagline, a "Deploy" button, and a yellow instruction bar that says "Enter details & select deficiency to get recommendations.". Below this is a large image of various fresh vegetables and fruits in a paper bag. At the bottom of the main area, the text "Flavor meets wellness!" is displayed.

# Specifications

**Front End :** Streamlit

**Back End :** Fast API

**Model** : K Nearest Neighbor (KNN)  
: Sentence Transformer  
(paraphrase-MiniLM-L6-v2)

**App** : [Culinary Compass](#)

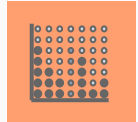
# Food Recommendation Using KNN



User Input:  
**Iron, Calcium,  
Niacin**



Recommends **top N food**  
items that best match the  
deficiency. **Sorghum,  
Fireweed, Tofu ...**

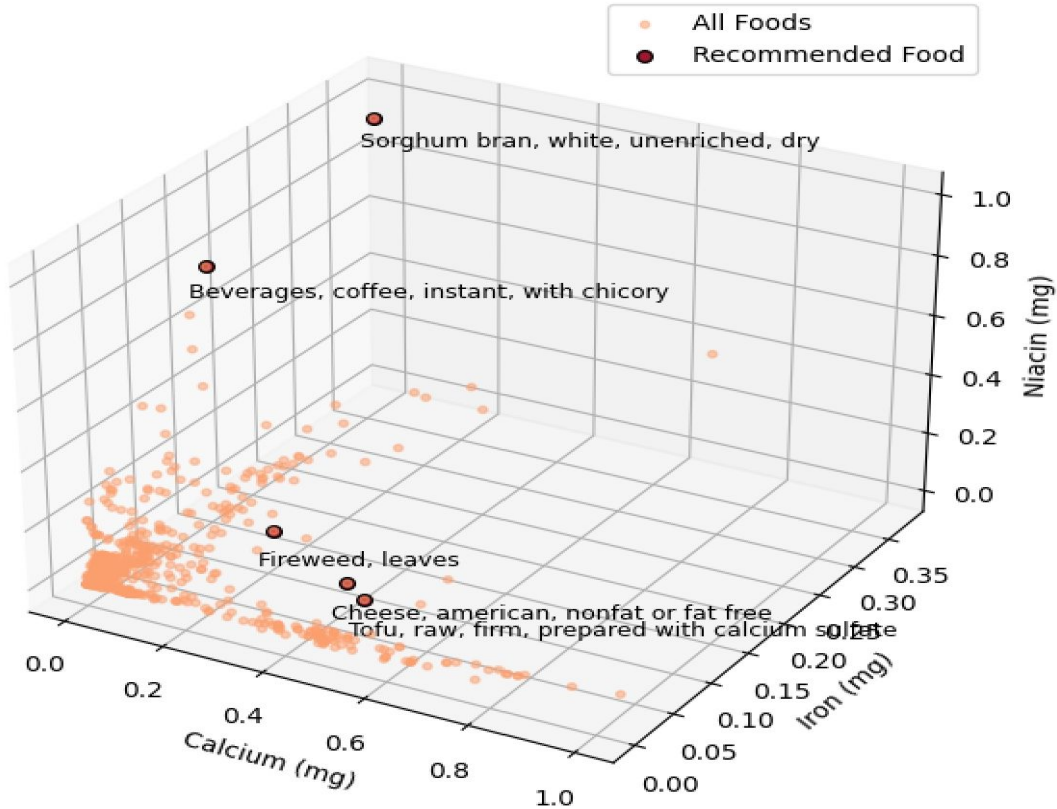


Create Binary Vector  
for Deficiencies  
Example:  
**[Iron: 1.0, Calcium: 1.0,  
Niacin:1.0, Zinc: 0.0,  
Sodium:0.0 ...]**

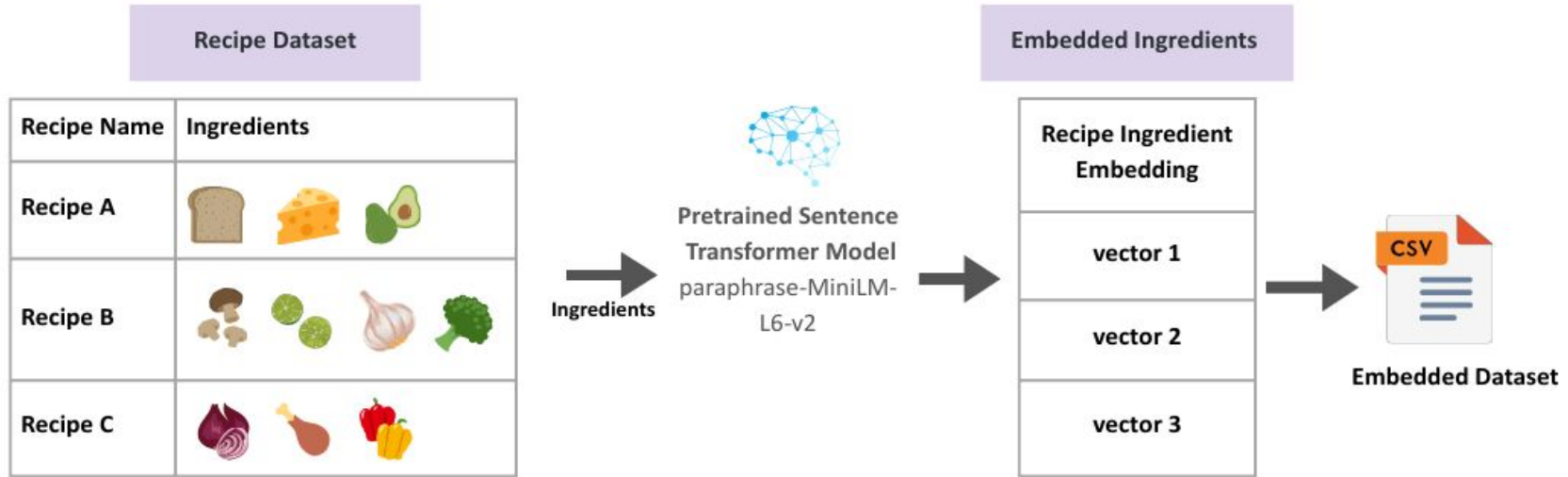


Apply **KNN**  
Calculate **Euclidean  
distance** between  
deficiency and nutrient  
values.

# KNN - Based Food Recommendation for Calcium, Iron, Niacin

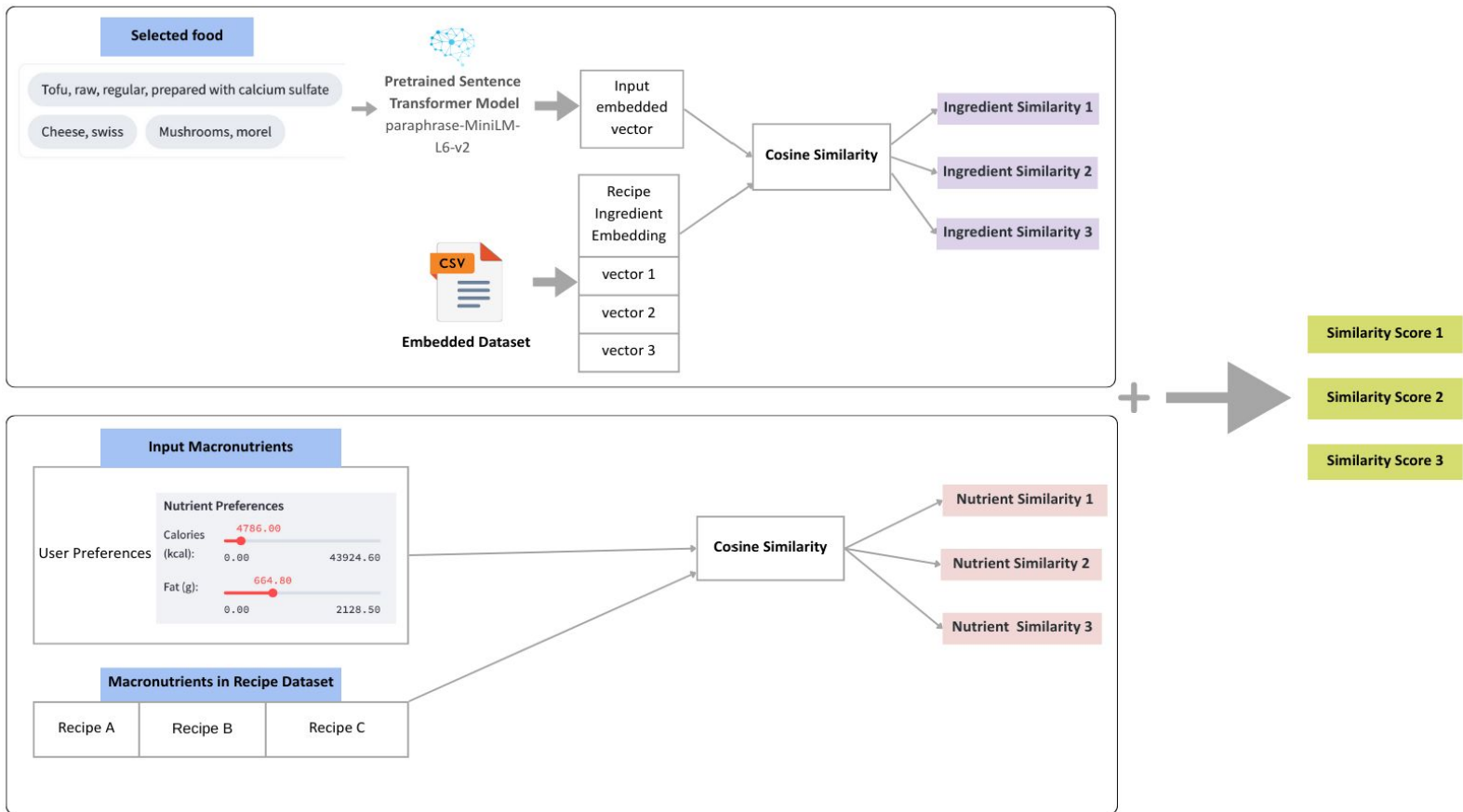


# Recipe Recommendation - NLP (Sentence Transformer)

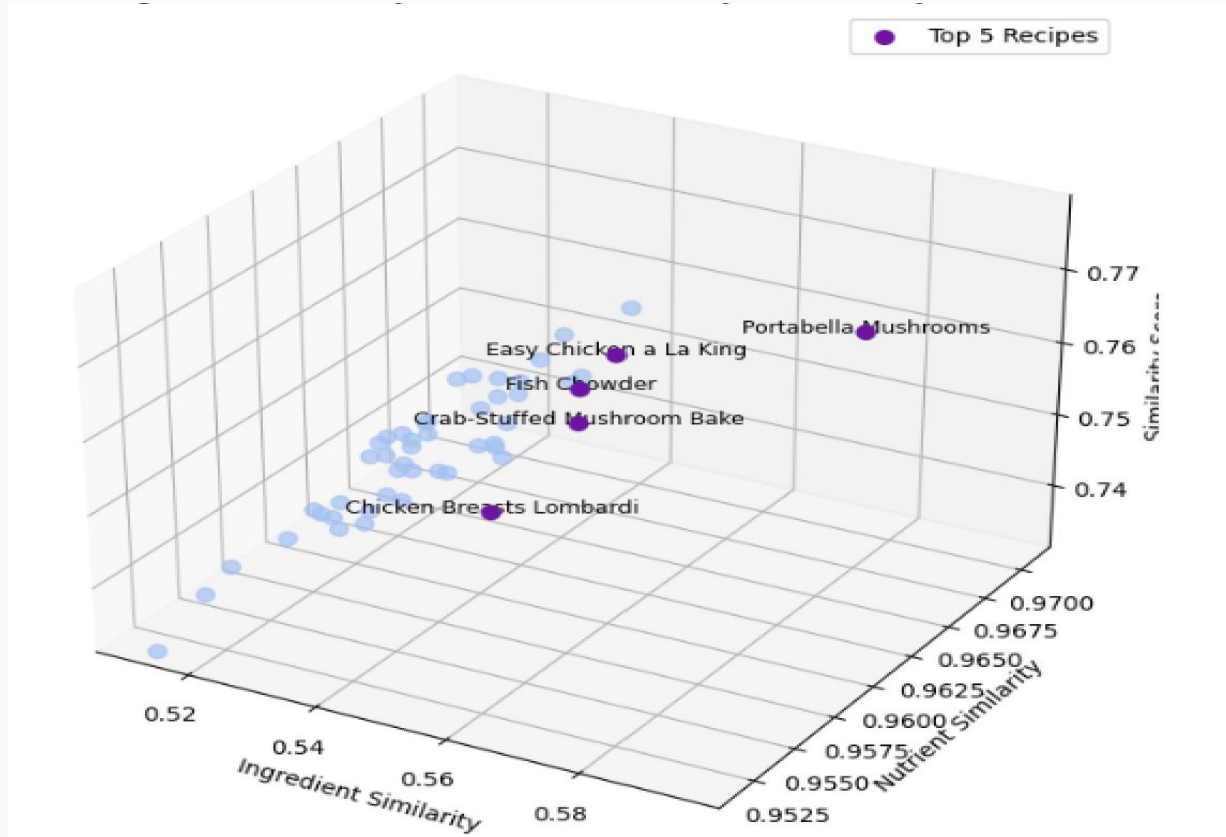


**Recipe Vectorization**

# Ingredient - Nutrient Similarity With Cosine Distance



# Ingredient Similarity vs Nutrient Similarity vs Final Similarity Score



# Tried and Tested

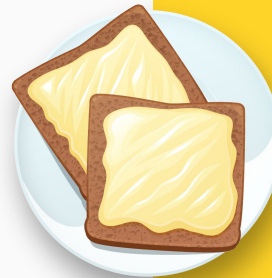
Food recommendation	Recipe Recommendation
KNN , Euclidean distance	Sentence Transformer all-MiniLM-L6-v2, Euclidean distance
KNN, Cosine similarity	Sentence Transformer all-mpnet-base-v2, Euclidean distance
DBScan	Sentence Transformer paraphrase-MiniLM-L6-v2, Cosine similarity





# Future Improvements

1. Optimize Embedding Generation for Large-Scale Datasets
2. Enhance Model Performance with Optimized Algorithms and Reduce Response Time
3. Integrate Collaborative Filtering to personalize recommendations based on user preferences



Any  
Questions?

Thanks !

