NutriSnap

Github: TBD
Authors: Vineet Jammalamadaka, Prakrit Baruah, William Belcher

Problem Statement (short)

The project aims to simplify the process of food tracking and nutritional analysis by replacing cumbersome manual data entry with a seamless, AI-powered system that accepts multi-modal input like photos and voice.

Background and Motivation

Whether it's patients with Irritable Bowel Disease (i.e., Crohn's and Ulcerative Colitis) who seek to mitigate their food-triggered symptoms or it's health-conscious gym-goers who aim to maximise their protein intake, understanding what and how much we eat is important for millions of people. Most apps that target these users and help them gain clarity around their food intake, macros, and portions are cumbersome, poorly designed, and ugly. Apps such as Chronometer require users to input every single raw ingredient of their meals and the gram amount that they ate. This is a bad user experience. Our goal is to make this process seamless.

Scope and Objectives

We make an AI-powered mobile app which makes food tracking and analytics simple.

1. **Simple input**: Instead of asking users to manually type all raw ingredients of their meal, we allow **multimodal input** (e.g., voice recordings, images, or text) of full meals.
2. **AI Extraction**: Our AI then extracts multi-modal input into relevant structured data. This will include information such as macros (e.g., protein intake) and potential symptom triggers (e.g., lactose).
3. **Useful analytics**; Based on extracted inputs, we can show useful trends over time and analytics. Examples include: amount of protein consumed daily, number of trigger foods eaten over time.
4. **Potential diet extension**: we can combine food tracking with many useful extensions for users. We can act as a user's diet companion where a user enters a specific diet (e.g., "paleo") and we can recommend which foods are in-line with this diet and how effectively someone is adhering.

Source of Data

Our primary data source will be a combination of publicly available datasets and data generated by the team for testing purposes. We plan to utilize image and text datasets related to food items and their nutritional content. For example, we will likely use datasets from Kaggle, such as the Food-101 dataset, which contains 101,000 images of 101 food categories, and the U.S. Department of Agriculture's FoodData Central for nutritional information.

The most relevant features will be the visual attributes of food (e.g., shape, color, texture), textual descriptions, and nutritional values (e.g., protein, fat, carbohydrates, and specific allergens or triggers like lactose).

Data Quality Concerns

- **Image variability**: The same food can look very different depending on lighting, portion size, and preparation. This can make accurate identification difficult.
- **Macro Identification**: Identifying macronutrient contents and weights/mass from images could pose a challenge.
- **Preliminary plan**: We will start with a small, curated dataset for initial model training and gradually expand.

Define Scope and Preliminary Design

We will develop a mobile application with the following core objectives:

- **Ingestion**: Allow users to submit meal data via photos or text descriptions of full meals, not just raw ingredients.
- **AI Processing**: Utilize a multi-modal model to identify food items and extract relevant structured data, such as macros and potential symptom triggers.
- **Personalized Analytics**: Display insightful trends and visualizations to help users understand their dietary habits over time.

Application Mock Design

Link to Figma Prototype

Research and Development

- Here is an article that discusses current food image recognition techniques and strategies, and is very up to date on current data sets.
- Here is another article, this one more specifically around volume estimation approaches.

- Here is a pubmed article about dietary assessment using multimodal inputs, noting how well multimodal models perform without significant outside training data sets.
- Here is an article discussing dietary trigger tracking in real time.

Limitations and Risks

- **AI Accuracy**: The primary risk is that our model may not always accurately identify foods from images, especially with complex or mixed-ingredient meals. Further, correctly identifying macronutrients from a picture could prove challenging.
- **Data Scarcity**: It may be challenging to find or create a sufficiently large and diverse dataset of full meals with accurate nutritional information.

Milestones

1. October 15: Train a preliminary model on a small, curated dataset of food images and develop a basic API for inference.
2. November 1: Build the core mobile app interface to allow photo and text input.
3. November 15: Integrate the AI model with the mobile app and test data extraction.
4. December 1: Develop the analytics dashboard and basic dietary guidance features.
5. December 10: Final touches and deliver demo at the showcase.

Note: Generative AI was used in the assembly of this statement of work, helping to source data sets and provide clarity and consistency checking after initial preparation.