

SmartEats: Your Personalized Calorie & Nutrition Advisor

Team Member: Jiayi Sun, Ninghui Hao, Qianwen Li, Taiyang Chen, Yantong Cui

Background:

In recent years, the global rise in nutrition-related chronic diseases such as diabetes, obesity, and cardiovascular diseases has become a significant public health concern^[1, 3]. These conditions are closely linked to poor dietary habits, emphasizing the need for a better understanding of meal nutritional content and the adoption of healthier eating practices. However, for most individuals, tracking daily food intake and assessing potential health risks can be a complicated and time-consuming task. This brings machine learning (ML) and natural language processing (NLP) tools into focus, offering the potential to automate and simplify this process.

By utilizing large datasets of food, nutrition, and health information available online, artificial intelligence (AI) models can help detect patterns in dietary habits that may increase the risk of developing nutrition-related diseases. Large language models (LLMs) provide users with personalized, real-time dietary suggestions based on their individual needs. For example, ChatDiet is a framework that uses LLMs for tailored nutrition recommendations, showcasing the potential of AI in this domain^[2]. However, there remains a gap in integrating deep learning (DL) models like convolutional neural networks (CNNs) to perform advanced image recognition, provide detailed nutritional analysis of daily meals and provide diet advice, which this project aims to address.

Problem Statement:

This project aims to develop SmartEats, an AI-powered application that analyzes food images, predicts disease risks based on dietary habits, and offers personalized nutrition advice. It will first require users to upload food pictures, and then it will use image recognition to estimate nutritional components like protein and fats and also display the amounts of each component ideally. Based on the results, SmartEats will calculate calories, assess potential health risks, and provide tailored dietary suggestions, such as healthy recipes.

SmartEats is designed for a wide audience, with a special focus on individuals with specific dietary needs, from those looking to lose weight or track daily calories to fitness enthusiasts aiming to optimize their nutrition intakes. Also, it is useful for individuals with a family history of diet-related chronic conditions, helping users monitor and manage their diet for better long-term health.

In this project, we intend to train the following models and integrate them into SmartEats:

- An image recognition model fine-tuned to food images to identify the food in images uploaded by users
- Tokenization tools for separating the components in food and linking them to nutrients and calories
- Logistic Regression or Random Forest models to predict disease risk based on nutrition data
- A LLM fine-tuned to professional nutrition Q&A which can provide users with real-time nutrition advice. We'll use ChatDiet as a source of inspiration and the API of other open-source LLMs.

To train our models, we require four types of data:

- Datasets containing diverse food images
 - [Food Image Classification Dataset](#): an image dataset containing 24K food images in 35 types of Indian and Western appetizers
 - [Fast Food Classification Dataset - V2](#): an image dataset containing images of 10 different types of fast food
- Datasets linking food to nutrition and calories (ideally including seasoning calories)
 - [FoodData Central](#): USDA's comprehensive source of food composition data with multiple distinct data types
 - [Food and their calories](#): food and their calories based on serving

- [Calories in Food Items](#): calories for 2,225 food items per 100 grams across 44 food categories
 - [FOODD](#): food detection dataset for calorie measurement using food images
- Datasets linking nutrition to disease
 - [NHANES](#): National Health and Nutrition Examination Survey data for nutrition-related health risk prediction
- A dataset containing nutrition advice, including dietary preferences and allergies
 - [DietaryChoiceAssistant](#): nutrition profession Q&A dataset

SmartEats will feature two main interfaces:

- A camera interface for capturing or inputting food images.
- A chatbot interface to answer the user's questions about nutrition advice.

Limitations and Risks:

- **Image Quality:** Ingredient identification and portion estimates may be affected by photo quality, including angles and lighting. To improve accuracy, users can manually input ingredient amounts, with defaults provided by model predictions.
- **Dataset Limitations:** SmartEats relies on diverse datasets to calculate calories, suggest recipes, and predict health risks. While we have key datasets, expanding coverage for uncommon foods and additional datasets may be needed in implementations and future updates.
- **Food Allergy Risk:** To minimize the risk of allergic reactions, we plan to include additional datasets of common allergens. SmartEats will flag recipes containing ingredients that may trigger allergies, alerting users to proceed with caution.
- **Computational Limitations:** Integrating large language models (LLMs) for personalized nutrition advice and interacting with users could face computational challenges, affecting real-time performance and scalability. These limitations may impact the application's responsiveness.
- **Privacy and Data Security:** Protecting user data is essential. Ensuring compliance with privacy laws like HIPAA and maintaining secure data storage will be crucial to building user trust and adoption.

Fun Factor:

SmartEats is like having a 24/7 personal calorie and nutrition advisor right in your pocket!

Milestones:

- Nutrition Data collection, preprocessing and integration: 09/26
- Models training and fine-tuning: 10/09
- Build container, pipelines and etc: 10/15 (milestone 2 deliverables: 10/18)
- Build functional CLI-pipeline with integration: 10/28 (milestone 3: 10/31)
- Develop a user-friendly interface and connect to backend functions/models: 11/13 (milestone 4 deliverables: 11/15)
- Integrate Chatbot within the interface: 12/01
- Final testing and deployment: 12/09 (final deliverables: 12/11)

References:

- [1] A. Żarnowski, M. Jankowski, and M. Gujski, "Public Awareness of Diet-Related Diseases and Dietary Risk Factors: A 2022 Nationwide Cross-Sectional Survey among Adults in Poland," *Nutrients*, vol. 14, no. 16, p. 3285, Aug. 2022, doi: [10.3390/nu14163285](https://doi.org/10.3390/nu14163285).
- [2] "ChatDiet: Empowering Personalized Nutrition-Oriented Food Recommender Chatbots through an LLM-Augmented Framework." Accessed: Sep. 19, 2024. [Online]. Available: <https://arxiv.org/html/2403.00781v1>
- [3] O. of the Commissioner, "Improving Nutrition to Turn the Tide on Diet-Related Chronic Disease," *FDA*, Aug. 2024, Accessed: Sep. 19, 2024. [Online]. Available: <https://www.fda.gov/news-events/fda-voices/improving-nutrition-turn-tide-diet-related-chronic-disease>