NORTHEASTERN UNIVERSITY

IE 7374 Project Proposal

REFRIGERATOR FOOD TRACKER WITH RECIPE RECOMMENDER SYSTEM

SMART REFRIGERATOR AI PROJECT

GROUP 12
NORTHEASTERN UNIVERSITY
By Ahantya Vempati, Chantelle Chan, Jensen Ho, Shyam Patel

Final Topic Area

We will be focusing on NLP text based GPT2 model for this project. This model works best with the **RecipeNLG** dataset that we have chosen. There was the idea of having vision based model feeding data into a text based model, but vision is beyond the scope of this project.

Description

According to the United Nations Environment Programme (2024), the world wastes over 1 billion meals' worth of edible food every day. To help address this issue, this project aims to implement a generative AI model designed to help households minimize food waste through intelligent inventory tracking, expiration alerts, and personalized meal recommendations.

Looking ahead, future innovations may involve the implementation of AI agent systems that extend beyond basic food tracking. These changes could make it possible for automated restocking, tailored nutrition monitoring, and even collaborating with composters that break down food waste at home. Ultimately, the refrigerator would evolve from a storage unit into a comprehensive, proactive tool for sustainable food management.

Problem Statement

Households in common day struggle with managing food waste, managing meals, and maintaining proper nutrition with the food they currently have on hand. Customers usually forget what items they brought in the refrigerators which in turn causes ingredients to be expired and in the end a waste of money. Not only that but spending time and creating healthy meals based on your ingredients can take a lot of time and be overwhelming.

Our project aims to fix this by developing an intelligent kitchen assistant that integrates real-time inventory tracking, expiration tracking, and personalized recipe generation. We are using generative AI, Retrieval Augmented Generation, Model fine-tuning, and transformer architectures. Through all this the assistant will create user-specific natural recipes, and track data on expiration of goods.

We aim to use our solution to fix issues with household efficiency by reducing the waste of goods and decreasing the amount of time for meal planning. In a world where we have inflation, health awareness, and sustainability issues our tool will drive smarter and adaptive decisions with food.

Background

In the recent design of modern appliances, developers have used emerging technology to assist with food waste in household. This is particularly applicable with utilization of IoT-enabled appliances and AI-driven systems. Brands like Samsung, LG, and Whirlpool have made smart refrigerators with features like internal cameras, barcode scanners, and inventory tracking systems to help users keep track of their food storage more effectively. These smart appliances can also connect to smartphone apps to interact with voice assistants, which improves the user experience and reduces waste through better inventory management (Basit, 2024).

In addition to hardware advancement, existing solutions have implemented analysis of customers' grocery shopping habits, dietary preferences, and leftover ingredients to come up with effective meal plans and recommend actions before food expires. Features such as real-time inventory tracking, AI-generated meal suggestions using items that are soon-to-expire, and notifications for food nearing expiration. Beyond food management, these solutions focus on making the best use of products that are already in the fridge to promote sustainability and reduce further food loss.

Plan of Action

Model Selection and Team Contribution

Responsible	Tasks
Chantelle	Preprocessing – Cleaning the dataset
Jensen	Coding GPT2 model and parameters (Large lab way and not API or hugging face, with tuning and prompt engineering for increasing accuracy)
TBD	Training – Wait and monitor - Takes times but not mental energy.
Ahantya	Architecture research (how GPT works – algorithms and how GPT connects with other components in the application) – Create application architecture diagram
Chantelle	Testing – Validation dataset and accuracy scoring. End-to-End. Confidence is close to the ground truth
Shyam	Code for interface with streamlit.io (Interface for App on top of Langchain with GPT) (Python)
Shyam	Langchain AI agents for prompts (wrapper for GPT) (Python) • No API call to LLM as an endpoint (avoid API calling for generation) • Not focused on accuracy. Architecture is more important.

	• Finetuning and prompt engineering ok • Minimal, only for enhancing output. • Check for hallucination
TBD	Videos recording at the end – Business, Technology, and Demo

Completed:

Completed Tasks
Share Github link with team: https://github.com/jasonhops/IE7374
Find dataset with ingredients and recipes https://www.kaggle.com/datasets/paultimothymooney/recipenlg

First, we need to preprocess the data by downloading **RecipeNLG from Kaggle** and filter by complete rows with fridge ingredients (e.g., fruit and vegetable). The dataset will be used to train a GPT model with ingredients as input and recipes as output. Second, use AI Agents tools that can fetch nutritional data, or find substitutes. Finally, if there is time use vision based model to scan refrigerator contents visually as input for text-based models.

For the front end, we can use the platform Streamlit.io. For Retrieval Augmented Generation (RAG) we can use Langchain graph. Double check what is in the fridge for generated recipes and use the **RecipeNLG trained GPT model** for generating recipes. To begin we will have a interface for inputting ingredients and asking customer/user for input (accept recipe or reject recipe). The goal is we have an application that works end-to-end as the final submission

We will use GPT2 based model and train with the RecipeNLG dataset. The reason is GPT2 is latest open source model that we can train on as GPT3 and above is not open source by OpenAI. Langchain is building AI agent coding framework with python. For storage technology we have the idea of database: DuckDB, Apache Parquet, SQLite, pandas plus flat file.

Dataset Description

We will use RecipeNLG data source from Kaggle for ingredient and recipe data from https://www.kaggle.com/datasets/paultimothymooney/recipenlg

With dataset RecipeNLG_dataset.csv that is (2.29 GB) with 1312871 recipes that has complete information. Here is the format of the data:

Title	String
Ingredients	String
Directions	String
Link	String

Source	String
NER	String
#	Integer

For food category we will narrow dataset to ingredients that are in the fridge: Fruit or vegetable in fridge because bakery items tend to be dry and not in fridge.

Research Question

During the project we will be able to answer the following questions:

How can fine-tuned GPT-2 models improve the quality and relevance of recipe generation using limited refrigerator inventory as input?

How does integrating a Retrieval Augmented Generation (RAG) approach impact the factual accuracy and nutritional value of AI-generated meal suggestions?

What are the key limitations of generative models in creating recipes with only partially complete ingredient lists, and how can prompt engineering mitigate these?

Expected Outcome

The smart refrigerator AI project is designed to deliver a functional AI solution for managing food inventory, reducing waste, and enhancing meal planning in a household or commercial kitchen. The initial implementation will focus on text-based inputs and with visual tracking added later. Both will be based on generative AI and agent-based architecture.

In the first phase, the text-based smart inventory will allow users to manually input food items, quantities, and expiration dates into the system. At mealtime, the system will generate personalized recipe recommendations based on current inventory and food nearing expiry. The system will update inventory when items are added, removed, and used. The smart refrigerator system supports sustainability by reducing food waste and encouraging efficient cooking.

Interface ideas include button to increment timeline, notification based on expiry and recipes as a suggestion. Recipes use decided by customer, only if they want the suggestion. Include box to talk to agent for testing. This is an end-to-end project that can show project completion. Different aspect can be talked about. Simple user interface is acceptable.

Our project will have the impact of reducing time and mental energy of managing food supply and actively combat food waste. Lastly, visual tracking will be added on top of text-based model in an AI Agent architecture. Inventory usage data can be used in pretraining, fine-tuning, and RAG workflow for refining recommendations.

Team and Schedule

Name	Role
Ahantya Vempati	Proofreader
Jensen Ho	Organizer
Chantelle Chan	Analyst
Shyam Patel	Manager

We will use the Scrum project management process to help achieve successful project delivery for Milestone 1, 2, 3, and Final Submission.

Module 9: Milestone 1: Final Project Proposal is due June 29

Module 11: Milestone 2: Data Pipeline is due July 13

Module 12: Milestone 3: Model Pipeline Implementation due July 20

Module 14 - Final Project Submission due August 3

Citation

United Nations Environment Programme. (2024, March 27). *World squanders over 1 billion meals a day: UN report*. https://www.unep.org/news-and-stories/press-release/world-squanders-over-1-billion-meals-day-un-report

Basit, Abdul. (2024, December 7). Smart Storage: How Technology is Revolutionizing Food Preservation. https://techbullion.com/smart-storage-how-technology-is-revolutionizing-food-preservation/

Plessers. (n.d.). *How smart refrigerators are changing food storage*. https://www.plessers.com/blog/how_smart_refrigerators_are_changing_food_storage