

Diana Montero & Joshua Moore

NextBite AI

1. Project Overview

- **Project Title:** AI-Powered Food Waste Reduction Using a Pre-trained LLM (Gemini) and RAG System to Help Waste Less and Transform Forgotten Food Scraps and Leftovers Into Delicious Meals

- **Problem Statement:** Food waste is at an all-time high. According to Rubicon, 30-40% of the food supply in the United States goes to waste. The annual food waste in America has an approximate value of \$161 billion. The average American family of four wastes around \$1,500 worth of food per year. Household food waste is both an environmental and financial problem. This will allow people to use food that would otherwise be wasted in new meals for themselves or their family. Solving this problem is important because it promotes sustainable living, saves money, and encourages a view where leftovers/food scraps can be turned into new foods rather than being viewed as trash.

- **Objectives and Goals:**

- Use a fine-tuned pre-trained LLM model (Gemini) to generate recipe names, steps, and helpful food storage tips.
- Integrate a RAG system to enhance responses with recipe data
- Build an application where users input 2-3 leftover ingredients, and it returns 1-2 beginner-friendly recipes, plus helpful food storage tips
- Success will be determined if the system outputs information with 80% accuracy

2. Scope and Use Case

- **Target Audience:**

- Families looking to save money and reduce waste
- Single household owners, college students, young professionals, and small families are trying to save money
- People who care about sustainability and want to minimize food waste

- **Use Case Scenarios:**

- A parent is trying to come up with a new dish for their family, only to find out they only have leftovers from previous meals. They don't have time to go to the store so they use this AI system to come up with something new out of the old.
- A college student in a dorm has limited cooking skills and a tight budget. They only have a few ingredients before payday. This AI tool will give them suggestions based on what they have, along with added storage tips.

- A young adult has just moved out of their parents' house and is learning how to cook. They have leftover ramen, mac and cheese, and half a tomato. They type in these foods and get an output "5-Minute _____" recipe with simple steps and a tip for extending tomato shelf life.
- **Project Constraints:**
 - Not every food combination will have a usable recipe
 - Not all food scraps or leftovers have well-documented recipe ideas in existing databases
 - Recipe suggestions may be limited to mainstream cuisines

3. Data Requirements

- **Data Sources:**
 - Kaggle datasets (for recipe data) (open-source)
 - USDA, FDA, NHS (for safe food storage guidelines) (open-source)
 - Food blogs, zero-waste recipe sites, and leftovers-focused cookbooks (proprietary sources)
- **Data Preprocessing Needs:**
 - Make ingredient names consistent
 - Deleting duplicates, incomplete recipes
 - Clean up the text
- **Data Labeling:**
 - Labeled data will be used from a mix of pre-existing labels, manual labeling, and automated techniques.
 - Existing recipe categories and ingredients
 - "Storage tip"

4. Model Selection

- **Model Type:**
 - Fine-tuned from a pre-trained LLM model while we feed in our data from the chosen resources
- **Model Justification:**
 - It is best for the problem at hand because it's efficient and allows for easy user access
 - The model will be readily available for any user to pull up and ask questions, with responses being readily available at the click of a button
 - Most efficient (saves time)
- **Transfer Learning (if applicable):**
 - Gemini

- Saves time using these models as they are already trained on large amounts of general text
- By fine-tuning it with a smaller dataset of recipes and food storage tips, it will teach the model to focus on the idea of this project, such as turning leftover ingredient lists into beginner-friendly recipes and safe food storage tips

5. Project Timeline and Milestones

• Timeline:

Week 1 - Figure out what the finalized AI tool will look like (Create a mock-up/outline)

Week 2 - Find Datasets and that can help feed our AI Model.

Week 3 - Clean and analyze data for model training and develop the website's interface

Week 4 - Connect the frontend with the model

Week 5 - Test the system, debug issues, refine outputs

Week 6 - Gather user feedback and improve the system's accuracy and functionality

Week 7 - Deploy to the cloud

Week 8 - Make presentation, finalize documentation

Week 9 - Practice presenting over and over

Milestone 1: Completing the Model itself: (Week 1-3)

☐ Key Deliverables

- Have a mock-up and outline for the AI model and website
- Have the AI model fed with proper datasets
- Be able to interact with the AI model from a user interface

Milestone 2: Complete the Website setup (Week 4-6)

☐ Key Deliverables

- Have the user interface properly implemented into a mock website
- Have the AI model respond properly to questions asked
- Making sure the API works with no bugs or glitches

Milestone 3: Make our Project Presentable (Week 7-9)

☐ Key Deliverables

- Have a fun and interesting design for our website
- Deploy website
- Test the website over and over
- Have a presentation and demo ready and heavily practiced

6. Evaluation and Testing Plan

- **Testing Environment:**

- Local development testing
- Postman for testing API endpoints
- Quality Assurance (output is grammatically correct and easy to understand, recipes match the given ingredients, and storage tips are relevant)

- **User Feedback Collection:**

- Do small user tests from different backgrounds to ensure the intuitiveness of the website's interface and AI model

- **Error Handling and Monitoring:**

- Regular monitoring will help us decide when to retrain or fine-tune the model if accuracy drops

7. Ethical and Responsible AI Considerations

- **Data Privacy:**

- Users won't be required to input sensitive data; therefore, no sensitive user data will be stored

- **Fairness and Bias:**

- Recipe data will most likely come from Western or mainstream cuisines, overshadowing cultural variety in suggestions
 - The plan would be to diversify the dataset by adding recipes from food blogs and cookbooks that focus on global cuisines

- **Environmental Impact:**

- The model we plan to use is relatively small, which uses less power compared to larger models
- Track carbon emissions during training
 - CodeCarbon (measures carbon emissions from Python scripts)

8. Project Outcomes and Impact

- **Expected Impact:**

- This project will help users reduce food waste, save money, and get creative in the kitchen by turning leftover ingredients into easy, beginner-friendly meals. It encourages sustainable habits, especially among students, young adults, and families who want to stretch their groceries and learn more about food storage. By making cooking less wasteful and more approachable, the tool supports both financial and environmental well-being.

- **Future Extensions:**

- Could offer recipe suggestions based on dietary restrictions or more global cuisines
- Add the ability for users to upload a photo of their ingredients
- Let users save favorite recipes

9. References

• Resources and Research:

- U.S. Food Waste Statistics Rubicon: <https://www.rubicon.com/blog/food-waste-facts/>
- USDA Food Loss and Waste Facts: <https://www.usda.gov/foodwaste/faqs>

• Data Sources:

- Food.com Recipes **(WILL USE)**(Kaggle):
<https://www.kaggle.com/datasets/irkaal/foodcom-recipes-and-reviews>
- Epicurious Recipes (Kaggle): <https://www.kaggle.com/datasets/hugodarwood/epirecipes>
- USDA FoodKeeper App: <https://www.foodsafety.gov/keep-food-safe/foodkeeper-app>
- FDA Food Storage & Safety Info: <https://www.fda.gov/food>
- FDA Are You Storing Food Safely?**(WILL USE)**:
<https://www.fda.gov/consumers/consumer-updates/are-you-storing-food-safely>
- USDA Leftovers and Food Safety:
<https://www.fsis.usda.gov/food-safety/safe-food-handling-and-preparation/food-safety-basics/leftovers-and-food-safety>
- [Nutrition.gov](https://www.nutrition.gov) Safe Food Storage:
<https://www.nutrition.gov/topics/food-safety/safe-food-storage>

10. High-Level AI System Plan

Backend and Frontend Main Features

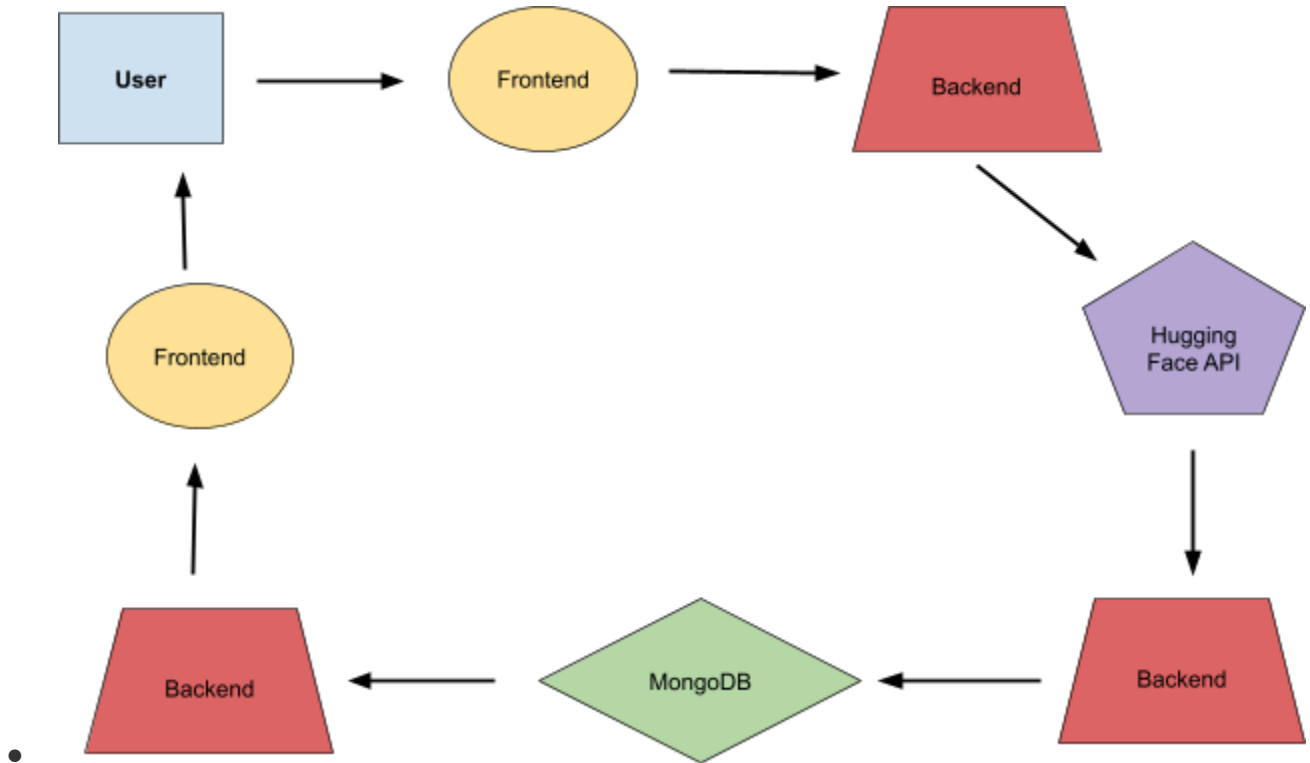
Frontend: React

- Input box
- Recipe result + tips section

Backend: Node.js + Express + MongoDB

- Uses AI to match ingredients with simple recipes and storage tips
- Keeps it fast, fun, and beginner-friendly

Connections



Inputs and Outputs

- Input: 2-3 leftover ingredients
- Output: 1-2 easy recipes with how to store leftovers safely

Ex:

- User types: "carrots, eggs, tortilla"
- AI replies:
 - "Try a veggie breakfast wrap!..."
 - "Or make carrot-egg fried rice!..."
 - "Store tortillas in the freezer to keep them fresh longer."

11. Model Prototype

- We started by importing the dataset path and printing the dataset in our program. Then we had the program separate each piece of data into the title of the recipe, ingredients of the recipe, directions of the recipe, and categories of the recipe. Then, we had the program drop any pieces of repeated ingredients and directions, while keeping the outputs in a lowercase font with no spaces in order to make reading more digestible. We then combined list into strings for model input,

we filtered only quick and easy recipes. Lastly, we encoded text into vectors using a pretrained sentence transformer (Roberta in Disguise).

- This model is pretrained on pretrained language data and transforms ingredient text into embeddings. (all-MiniLM-L6-v2 from sentence-transformers). We used the Gemini 2.5 flash model to generate a recipe. Data split 80% training, 10% validation, and 10% testing. We trained embeddings on the training set only.
- The evaluation metric we used was the cosine_similarity. It measures how similar the AI recipe is to the closest real one.
- When we asked for a recipe using the simple ingredients of chicken, rice, and onions, the output was a recipe for one-pot Speedy Chicken and Rice. A recipe yielding 2 servings and taking about 35 minutes to complete. It gave ingredients, equipment needed, and the instructions were simple to follow. From chopping chicken and sauteing the onions. To boil the rice and use the right amount of seasoning. Mixing to create the perfect simple recipe at home for anyone to cook on a budget. It even gave a cute little message, "Enjoy your delicious, homemade, and super easy chicken and rice!"
- We only used one metric. The plan is to add more evaluation metrics like BLEU or ROUGE. We need to do an evaluation on validation/ test sets.