Meal And Recipe Generator (MARG)

AI-Powered Meal Generation System Proposal

Team: Daniel Mehta and Thomas Nash

1. Problem Statement

Many people face the daily challenge of deciding what to cook with the available ingredients in their kitchen. This challenge leads to:

- Food waste occurs when ingredients aren't used before spoiling
- Unnecessary grocery purchases when existing ingredients could be combined effectively
- Missed opportunities for creative meal preparation
- Time wasted searching through recipe websites with incomplete ingredient matches

This problem is ideal for a generative AI solution because:

- Recipe creation follows patterns that can be learned from existing recipes
- Ingredient combinations and cooking techniques have underlying rules and relationships
- The solution requires both understanding ingredient compatibility and generating coherent, detailed instructions
- Traditional rule-based systems cannot capture the nuanced relationships between ingredients and cooking methods

2. Proposed Solution

We propose developing an AI-powered recipe generation system that creates detailed, customized recipes based on available ingredients.

Primary Components:

1. Ingredient Analysis Module
   - Processes user input of available ingredients and quantities
   - Identifies potential ingredient combinations
   - Suggests additional ingredients if needed
   - Optionally accepts tags such as dietary restrictions or themes (e.g., “gluten-free,” “spicy”)
2. Recipe Generation Module
   - Uses the Ashikan/dut-recipe-generator model
   - Creates complete recipes using available ingredients
   - Generates step-by-step cooking instructions
   - Provides cooking times, temperatures, and serving suggestions
   - Can reroll new recipes using the same input.

Model Selection:

- Primary: Ashikan/dut-recipe-generator
  - A BLOOM560 m-based model fine-tuned on the DUT Diabetic-Friendly Recipes and a cleaned version of RecipeNLG. It generates structured, coherent recipes with metric units and health-conscious guidelines.
- Optional: Stable Diffusion for recipe image generation

The generative aspect enables:

- Dynamic recipe creation based on specific ingredient combinations
- Adaptation of cooking instructions to available equipment
- Scaling recipes based on ingredient quantities
- Generation of coherent, detailed cooking procedures

3. Dataset(s)

We do not train our own model, so we rely on the datasets used to fine-tune the Ashikan model:

- **RecipeNLG**: A large dataset of structured recipes used in the initial training
- **DUT Diabetic-Friendly Recipes:** Custom-curated recipes reviewed by food science experts for nutritional value and clarity

We may optionally reference:

- **USDA FoodData Central**: For basic nutritional estimates
- **Public recipe datasets** (e.g., Food.com) for UI examples or tag matching logic

Preprocessing Requirements:

- Standardization of ingredient names and measurements
- Extraction of cooking techniques and times
- Classification of cuisine types and dietary restrictions (rule-based)
- Standardize user input (e.g., lowercase, remove typos)
- Format the prompt as a JSON object for the model

Ethical Considerations:

- Cultural sensitivity in recipe modifications
- Dietary restriction accuracy
- Attribution of recipe sources
- Bias monitoring in ingredient combinations
- Privacy protection for user data

4. System Design Overview

Architecture Components:

1. Input Processing Layer
    - Ingredient parser and normalizer
    - Quantity converter
    - Equipment availability checker
2. Recipe Generation Engine
    - Ingredient compatibility analyzer
    - Recipe structure generator
    - Instruction sequence creator
    - Scaling calculator
3. Output Generation Layer
    - Recipe formatter
    - Nutritional calculator
    - Alternative suggestion generator
    - Safety check system

Implementation Approach:

- Python-based development using modern NLP libraries
- API-first design for easy integration

- Open Source LLMs
- StreamLit for UI

Evaluation Metrics:

- Recipe success rate
- Ingredient utilization efficiency
- User satisfaction ratings
- Nutritional balance
- Instruction clarity
- Safety compliance

5. Challenges and Limitations

Technical Challenges:

1. Recipe Scaling
   - Solution: Create robust quantity adjustment algorithms
   - Validate scaling logic with expert chefs
2. Safety Considerations
   - Solution: Implement multiple validation layers
   - Include warning systems for potentially unsafe combinations

Timeline (6-Week Project):

Timing	Task Descriptions
Week 1	• Dataset collection and preprocessing • Initial system architecture setup • Basic ingredient parsing implementation
Week 2	• Build prompt formatting and testing with the model • Add dietary tag parsing and validation • Polish UI and reroll logic • Optional: nutritional info estimation
Week 3	• Instruction generation system development • Safety validation implementation • Initial user interface development
Week 4	• Recipe scaling system implementation • Integration testing • User feedback collection
Week 5	• System optimization • Error handling improvement • Documentation development
Week 6	• Final testing and validation • Performance optimization • Deployment preparation

6. References

1. Marin, J., et al. (2019). "Recipe1M+: A Dataset for Learning Cross-Modal Embeddings for Cooking Recipes and Food Images." IEEE Transactions on Pattern Analysis and Machine Intelligence.

2. Salvador, A., et al. (2017). "Learning Cross-modal Embeddings for Cooking Recipes and Food Images." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

3. USDA. (2019). "USDA Food Composition Databases." U.S. Department of Agriculture, Agricultural Research Service.

4. Brown, T., et al. (2020). "Language Models are Few-Shot Learners." arXiv preprint arXiv:2005.14165.

5. Prof Ashika Naicker*, Mr Shaylin Chetty, Ms Riashnie Thaver*, Ms. Anjellah Reddy*, Dr. Evonne Shanita Singh*, Dr. Imana Pal*, Dr. Lisebo Mothepu*. *Durban University of Technology, Faculty of Applied Sciences, Department of Food and Nutrition, Durban, South Africa

6. Bień, M., Gilski, M., Maciejewska, M., Taisner, W., Wiśniewski, D., & Lawrynowicz, A. (2020, December). RecipeNLG: A cooking recipes dataset for semi-structured text generation. In Proceedings of the 13th International Conference on Natural Language Generation (pp. 22–28). Association for Computational Linguistics. https://www.aclweb.org/anthology/2020.inlg-1.4

7. Naicker, A., Chetty, S., Thaver, R., Reddy, A., Singh, E. S., Pal, I., & Mothepu, L. (n.d.). *Durban University of Technology, Faculty of Applied Sciences, Department of Food and Nutrition, Durban, South Africa*.

8. *Fooddata Central Frontpage*. USDA FoodData Central. (n.d.). https://fdc.nal.usda.gov/

9. Streamlit Inc. (n.d.). *Streamlit* [Computer software]. https://streamlit.io/