Customized AI Kitchen For India

Rahul Krishnan,Abia Abraham,Belwin Vaniyapurayil Binoj,
Savio Saju,Ashique Muhammed
Mentor : Mrs.Starlet Ben Alex

Saintgits Group of Institutions

July 11, 2024

Acknowledgement

We would like to thank Intel for giving us this opportunity to work on this extremely innovative problem statement. We also want to express our sincere gratitude to our mentor Mrs Starlet Ben Alex Ma'am for her constant support and encouragement. We are especially thankful to our institution ,Saintgits College of Engineering,for providing the resources and environment conducive to research and learning.
Last but not the least we thank Siju Swamy sir for being the constant piller of support and guidance .

Introduction

CUSTOMISED AI KITCHEN:In today's fast-paced world, cooking nutritious and delicious meals can often seem like a daunting task. Busy schedules, diverse dietary preferences, and the challenge of making the most of available ingredients can make meal planning overwhelming. This is where our innovative project, the Smart Recipe Generator, comes into play.This project leverages the power of generative AI to create a unique and interactive tool for culinary enthusiasts. By training a language model on a dataset of popular Indian recipes, this project aims to generate new and customized Indian recipes based on user inputs. The tool is designed to assist users in discovering new dishes and experimenting with various ingredients and cooking techniques, all within the rich and diverse context of Indian cuisine.

Methodology

The methodology for creating and using an Indian recipe dataset to train an AI model involves data collection, preprocessing, model training, and deployment. This streamlined process ensures the dataset is comprehensive, high-quality, and effective for training a generative model like GPT-2.

- **Step 1**: Data Collection
  1. Sources: Collect recipes from online recipe websites, cooking blogs, cookbooks, and user-contributed platforms.
  2. Content:Include recipe titles, ingredient lists, and preparation steps.Our dataset consists of 1000 rows 3 coloumns.
  3. Diversity: Ensure the dataset represents various Indian regions and includes traditional and modern dishes.
- **Step 2**: Data Preprocessing
  1. Data Cleaning: Remove duplicates and irrelevant entries. Standardize measurements and ingredient names. Correct spelling errors and inconsistencies.
  2. Text Normalization: Convert text to lowercase.Tokenize text into words or subwords. Remove special characters and excessive whitespace.

Methodology

- **Step 3**: Model Training Dataset Preparation:
  Combine title, ingredients, and instructions into a single text column. Split the dataset into training, validation, and test sets.Tokenize the text data using the GPT-2 tokenizer. Pad and truncate sequences to a fixed length.
  Fine-Tuning GPT-2: Load the pre-trained GPT-2 model and tokenizer. Define training parameters (learning rate, batch size, epochs). Fine-tune the model on the recipe dataset using the GPT2LMHeadModel. Monitor training with validation loss and adjust hyperparameters as needed.
- **Step 4**: Model Deployment and User Interface Loading the Model: Load the fine-tuned GPT-2 model and tokenizer for inference. Creating the User Interface: Use Gradio to create an interactive web interface. Allow users to input ingredients or prompts and receive generated recipes.

Challenges in RECIPE GENERATOR

- **Data Collection and Quality**:
  - Diverse Data Sources: Aggregating recipes from various sources can be challenging due to inconsistent formats and varying levels of detail.
  - Data Accuracy: Ensuring the accuracy of ingredient lists, cooking instructions, and nutritional information is critical for user trust and satisfaction.
- **User Personalization**:
  - Preference Learning: Accurately learning and adapting to individual user preferences requires sophisticated machine learning algorithms and extensive user interaction data.
  - Dietary Restrictions: Handling complex dietary restrictions and preferences (e.g., allergies, specific diets) in a way that still provides a wide range of suitable recipes can be challenging.

Application of Indian Recipe Dataset

Indian recipe datasets have been utilized in various AI applications, including:

- **Recipe Generation**: Fine-tuning language models like GPT-2 to generate novel Indian recipes based on input ingredients or partial instructions. These models learn to mimic the style and structure of Indian culinary texts.
- **Ingredient Substitution**: Developing systems to suggest ingredient alternatives in Indian recipes, considering availability, dietary restrictions, and flavor compatibility.
- **Culinary Exploration**: Creating tools that allow users to discover new Indian dishes, explore regional variations, and experiment with different cooking techniques.

Results and Discussions

The project aimed to develop an AI-based recipe generator using a fine-tuned GPT-2 model trained on 1000 Indian recipes. Key outcomes include:

- **Model Performance**:The GPT-2 model was fine-tuned with low validation loss, indicating effective learning. The model generated coherent and contextually relevant recipes.
- **Coherence**: Logical sequences of ingredients and instructions were maintained.
- **Creativity**: The model introduced novel variations of traditional dishes. Relevance: Recipes reflected Indian culinary traditions accurately.
- **User Interaction**: An intuitive Gradio interface allowed users to input prompts and receive customized recipes.

Results and Discussions

- **Challenges and Limitations**
  - Data Quality: Minor inconsistencies in ingredient names and measurements were observed.
  - Cultural Accuracy: Some recipes included unusual ingredient combinations.
  - Recipe Completeness: Occasionally, generated recipes had missing steps or unclear instructions.
- **Future Work**
  - Dataset Expansion: Include more diverse and region-specific recipes, along with additional metadata.
  - Model Refinement: Further fine-tuning and incorporating multimodal data could improve performance.
  - User Feedback Integration: Continuously collect and incorporate user feedback for refinement.

Recipe Generator

Conclusion

- Indian recipe datasets play a crucial role in training AI models for culinary applications. By ensuring diversity, quality, and contextual richness, these datasets enable the development of advanced AI systems that can generate, recommend, and personalize Indian recipes. Addressing challenges in data collection, quality, and representation will enhance the effectiveness of these datasets and contribute to the broader field of AI-driven culinary innovation.

REFERENCE

- https://github.com/vishalmysore/cookGPT
- https://medium.com/@visrow/unleashing-the-flavor-cookgpt-an-ai-chef-for-indian-cuisine-af9369a1fe90