NutriBot: A Discord bot for personalized recipe recommendations and nutritional analysis using USDA and Recipe1M+ dataset

Abiral Shrestha

Jyoti Bhandari

Sudip Adhikari

Statement of Project Objectives

- Develop a recommendation system that can generate new recipe ideas based on users' dietary preferences and restrictions using pre-trained BERT model and Recipe1M+ and USDA National Nutrient Database datasets.
- Implement the recommendation system as a Discord bot that can interact with users and provide them with recipe suggestions along with nutritional information.

Statement of Value

- Accurate and Personalized Recommendations: Our project aims to provide accurate and personalized recommendations for recipes based on the user's dietary preferences, nutritional requirements, and ingredients available to them.
- Efficient and User-friendly Interface: Our system offers an efficient and user-friendly interface for users to easily obtain recipe recommendations, nutrient information, and cooking instructions with just a few clicks, saving them time and effort.
- Scalability: Our project can be scaled up to handle a large volume of user requests and data on nutrition and recipe preferences, making it useful for a wide range of users.

Review of the State of the Art and Relevant Works

Can a Chatbot Determine My Diet?: Addressing Challenges of Chatbot Application for Meal Recommendation

Poor nutrition harms immunity, development, and productivity. Chatbots can be virtual coaches for healthy eating but building them has challenges. This paper explains these challenges, covering technical, theoretical, behavioral, and social aspects. The paper proposes a pipeline as a guideline for developers to create robust chatbots.

FoodKG: A Semantics-Driven Knowledge Graph for Food Recommendation

A Comprehensive Knowledge Graph for Food Representation and Applications" by Zhang et al. This project aims to build a comprehensive knowledge graph for the food domain, which includes information on food items, nutrients, cooking methods, and more. The knowledge graph is constructed using a combination of manual curation and automated extraction from various data sources. It can be used for a variety of applications, including recipe recommendation systems, dietary analysis, and food-related chatbots.

Review of the State of the Art and Relevant Works

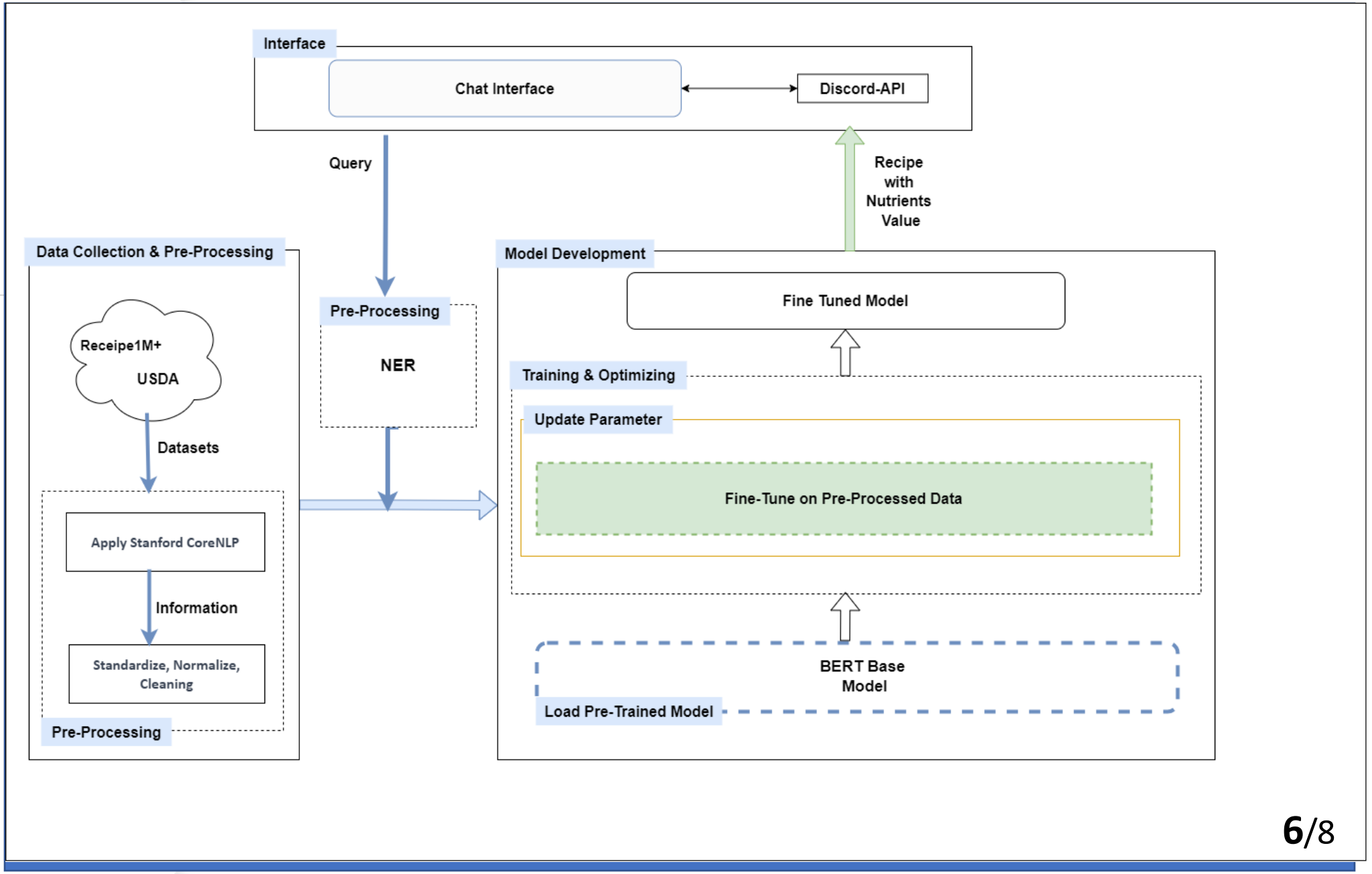
SlimMe, a Chatbot With Artificial Empathy for Personal Weight Management: System Design and Finding

Obesity is a growing issue, and AI technology provides an opportunity for dietitians to support weight loss with chatbots. This study created a chatbot called "SlimMe" with artificial empathy for motivational support. The system design of SlimMe involves the use of machine learning algorithms to predict user preferences and personalize recommendations for calorie intake. The chatbot also offers personalized meal planning and exercise suggestions to help users achieve their weight loss goals. The chatbot accurately responded to relevant user requests 67.38% of the time and did not respond 1.12% of the time.

Approach : Tools and Models

- Data Collection: Recipe1M+ dataset and USDA(United States Department of Agriculture.) dataset will be used to collect data on recipes, nutrition information, and ingredients.
- Data Preprocessing: The collected data will be preprocessed using techniques such as tokenization, stemming, and lemmatization to prepare it for further processing.
- Named Entity Recognition (NER), Part-of-Speech (POS) Tagging, and Dependency Parsing: Stanford CoreNLP will be used to perform NER, POS tagging, and dependency parsing on the preprocessed text data to extract relevant information such as ingredient names, measurements, and actions.
- Pretrained BERT Model: A pretrained BERT model will be fine-tuned on Recipe1M+ dataset and USDA dataset to develop a recipe recommendation system that provides personalized recommendations based on user preferences, dietary restrictions, and fitness goals.
- USDA National Nutrient Database – Dataset consisting of ID FoodGroup, Descrip, Energy_kcal, Protein_g, Fat_g, Carb_g, Sugar_g, Fiber_g, VitA_mcg, etc.
- Recipe1M Dataset (Layers and Ingredient detections files)
  - Layer1 – Dataset consisting of ingredients, url, partition, title, id, instructions
  - Ingredient detections – Dataset consisting of valid, id, ingredients
- **Discord API** for the chatbot integration.

Approach : Architecture



Deliverables

- A discord bot that provides personalized recipe according to the user's request
  - Input: Accept user queries through the user-friendly interface.
  - Output: Provide personalized recipe
- A project report, composed as a paper written in the style of an ACL/NeurIPS/AAAI, etc. Conference.
- A short video (< 5 minutes) with the demonstration of our project.
- A new Github repository with completed project (including. code and slides), along with a user documentation manual as a .MD file describing the project and usage instructions to other interested students and researchers.

Evaluation Methodology

- The model's performance will be evaluated based on its accuracy and F1-Score.
- The accuracy measures the overall correctness of the model's predictions i.e., the number of correct predictions made by the model divided by the total number of predictions made.
- F1-Score is calculated as the harmonic mean of precision and recall, where precision is the ratio of true positives to the sum of true positives and false positives, and recall is the ratio of true positives to the sum of true positives and false negatives.
- The model should have an accuracy greater than 67.38% (accuracy of state of the art "SlimMe").