## Introduction and Overview

**Overview**

This project outlines the development of a personalized diet recommendation system. The system will analyze patient data including weight, height, blood pressure (BP), thyroid or diabetic conditions, and cholesterol levels, along with their food preferences and weight goals. By leveraging this information, the system will categorize patients based on their needs and recommend a customized diet plan.

The recommendation process will involve two key components:

1. **Calorie Intake and Food Restrictions:** Based on the patient's profile, the system will determine their daily calorie needs and suggest foods they should avoid due to potential health concerns.
2. **LLM (RAG-based) Generation with Knowledge Graph**: The system will utilize a large language model (LLM) with Retrieval-Augmented Generation (RAG) to create personalized diet plans. Additionally, a knowledge graph will be used to inform the LLM, ensuring recommendations are tailored to individual dietary needs and preferences.

**Purpose**

This document serves as a comprehensive project implementation plan, detailing the development stages, functionalities, and goals of the diet recommendation system.

**Scope and Objectives**

The scope of this project encompasses the design, development, and implementation of the diet recommendation system.

The primary objectives are:

1. To develop a system that analyzes patient data and food preferences to recommend personalized diet plans.
2. To integrate RAG-based LLM generation to create diverse and adaptable meal recommendations.
3. To leverage a knowledge graph to enhance the accuracy and personalization of dietary suggestions.
4. To present a user-friendly interface for accessing and managing personalized diet plans.

## Implementation Strategy

**Methodology**

Data preprocessing, Knowledge Graph (KG) Population and RAG System Development

1. **Data Collection and Preprocessing:** Collect patient data (weight, height, BP, medical conditions, cholesterol) and food preferences. Preprocess the data for consistency and prepare it for knowledge graph ingestion.
2. **Knowledge Graph (KG) Population:** Utilize an LLM to extract relevant information from patient reports to design Knowledge Graph Schema and populate the knowledge graph with entities (patients, foods, conditions) and relationships (patient weight, food calories, condition limitations).
3. **RAG-based System Development:** Develop a Retrieval-Augmented Generation (RAG) system within the LLM. This system will retrieve relevant information from the knowledge graph to inform personalized diet suggestions.
4. **Expert Moderation:** Implement a UI for domain experts to review and modify LLM-generated suggestions based on their knowledge.
5. **Data Storage:** Store original patient data, processed data, LLM-generated suggestions, and expert-moderated data in a secure database.

## System Architecture

## 

## 

## 

## 

## Data Management

**Data Entities:**

Diet Plan: This represents a complete dietary plan for a specific user.

* Goal: Weight loss or weight gain
* Cuisine: Regional cuisine preference (South Indian, North Indian, Maharashtrian)
* Dietary Restrictions: Vegetarian (Veg) or Non-vegetarian (Non Veg)
* Medical Conditions: Presence of Thyroid or Diabetes
* Total Calories, Proteins, Carbs, Fats, and Fibers: Overall nutritional content of the plan
* Meal: Breakfast, Lunch, Dinner, Pre-workout snack, Post-workout snack, Morning snack, Evening snack
* Meal Items: Comma-separated list of food items included in the meal (e.g., Upma, Tomato peanut chutney)
* Serving Size: Not explicitly provided in this sample

User Data:

* BMI (Body Mass Index)
* Body Fat Percentage
* Glucose Levels (potentially for Diabetes management)
* Height and Weight
* Blood Lipid Profile: (Tc/HDL, LDL/TG)
* Blood Pressure (Diastolic and Systolic)
* HbA1c (Glycated Hemoglobin - potentially for Diabetes management)
* Age
* Gender
* Food preference : veg, non veg
* Food type: south indian, north indian

**Data Preprocessing:**

Available tools will be utilized to clean and preprocess the dataset to make it suitable to be fed into a knowledge graph.

**Data Storage:**

A knowledge graph database can store information about diet plans (goal, cuisine, restrictions, conditions), meals (items), and potentially user data (if collected directly).

A separate KG would link diet plans to specific user profiles based on their medical conditions and goals.

**Additional Considerations:**

1. Nutritional information for individual food items is not provided.
2. Gender, Age of individuals is not known.
3. Amounts of consumption of food items also needed to be recorded.
4. Personalization Techniques:

* Incorporating User Preferences: The system prioritizes user-specified cuisines and dietary restrictions when recommending meals.
* Condition-Specific Recommendations: Based on the user's medical conditions (accessed from the KG), the LLM suggests foods that promote overall health (e.g., high-fiber foods for Diabetes).
* Calorie and Nutrient Balancing: The system calculates ideal calorie intake based on user data and creates a balanced diet plan meeting recommended macronutrient (carbs, protein, fat) and micronutrient (vitamins, minerals) goals.

## Recommendation Method

This system utilizes a combination of Large Language Models (LLMs), a knowledge graph (KG), and user data to generate personalized diet plans:

**Prompt Engineering with and Knowledge Graph Population**

**Prompt Engineering:**

We will utilize prompt engineering techniques to make sure we extract the relevant information from the provided reports and user profiles to build a knowledge graph.

**Knowledge Graph Population:**

Construct a KG to represent entities and relationships relevant to diet recommendations.

Example:

Entities:

Foods, nutrients, medical conditions, cuisines, user profiles

Relationships:

"HasNutrient" (Food - Nutrient, e.g., Upma - Carbohydrates)

"Increases/Decreases" (Nutrient - Medical Condition, e.g., Fiber - Lowers Cholesterol)

"Prefers" (User - Cuisine, e.g., John - South Indian)

"Manages" (Diet - Medical Condition, e.g., Weight Loss Plan - Diabetes)

**Personalized Diet Plan Generation**

**User Input and Profile Creation:**

Goal: Weight loss or weight gain

Dietary Restrictions: Vegetarian, Non-vegetarian, Allergies

Medical Conditions: Diabetes, Thyroid (if applicable)

Cuisine Preferences: South Indian, North Indian, etc.

Additional Information: Weight, Height, Activity Level (for calorie calculation)

**Personalized Recommendation:**

The system leverages the LLM and KG to generate a customized diet plan.

The LLM, trained on dietary data, can suggest appropriate foods and meals based on user preferences and conditions.

The KG provides context and relationships between entities (e.g., suggesting low-carb foods for Diabetes management).

User data is used to personalize calorie intake and portion sizes.

To assess the effectiveness of the recommendations, consider these metrics:

* User Satisfaction: Surveys or feedback mechanisms to gauge user experience and satisfaction with the generated diet plans.
* Nutritional Adequacy: Evaluate if the recommended plans meet recommended daily allowances (RDA) for macronutrients and micronutrients.
* Adherence Rates: Track how well users follow the recommended plans over time.
* Health Outcome Improvement (Long-term): Monitor changes in user health markers (weight, blood sugar levels) over a longer period

## Features and Functions for the User Interface

**User Input and Profile Creation**

* Input fields for user goals (weight loss/gain)
* Selection options for dietary restrictions (vegetarian, non-vegetarian, allergies)
* Ability to indicate medical conditions (diabetes, thyroid)
* Cuisine preferences (South Indian, North Indian, etc.)
* Fields for weight, height, and activity level

**Profile Management**

View and update user profile information (weight, and height needs to be updated timely)

**Diet Plan Access and Management**

* View personalized diet plans generated by the system
* Ability to see breakdowns by meal (breakfast, lunch, dinner, snacks) with specific food items
* Functionality may be limited to viewing recommended foods without quantities for initial release

**Timeline/Schedule**

| **Phase** | **Task** | **Subtask** | **Estimated Duration**  **(6 Weeks)** |
| --- | --- | --- | --- |
| Phase 1: Knowledge Graph Population & RAG System | Data Acquisition and Preprocessing | Data collection | 1 week |
|  |  | Develop data cleaning protocols (handling missing values, inconsistencies) |
|  |  | Format data for KG ingestion (standardization, mapping to KG entities) |
|  | Knowledge Graph Population | Design KG schema (refine entities) | 1 week |
|  |  | Configure LLM for information extraction (training on medical reports, food data) |
|  |  | Populate KG with patient data (extracting relevant information from reports) |
|  | RAG System Development | Define retrieval strategies (keywords, entity relationships) | 1 week |
|  |  | Integrate KG access within LLM (enabling retrieval of relevant information) |
|  |  | Implement logic for suggesting diets (based on retrieved information and user preferences) |
|  | Expert Moderation | Design user interface for expert review (presenting LLM suggestions and patient data) | 2 weeks |
|  |  | Implement functionalities for modification and approval (allowing experts to adjust suggestions) |
|  | Data Storage | Select and configure database system (secure storage for patient data, KG, and recommendations) | 1 week |
|  |  | Secure data storage protocols (encryption, access control) |