**DIET MODULE:**

Three datasets were developed in this project to implement diet recommendation. One for breakfast, one for lunch and dinner and one for snacks. The breakfast dataset is further divided into multiple sections like the main meal, milk, nuts and fruits. The lunch and dinner dataset is divided into multiple sections like the main meal, bread/rice and salads. The snacks dataset is divided into main meal, fruits and sweets. The dataset is created with care under the guidance of a certified Nutritionist- Dt. Dhwani Parekh to consider if a food item is suitable for a user with various health issues like diabetes, thyroid, PCOS, kidney disease and lactose intolerance. Furthermore, the dataset incorporates user preferences like veg, non-veg or vegan, north Indian and south Indian for cuisine preferences.

1. **Sizes of different datasets:**

**Breakfast:**

    Main Meal - 23

    Fruits - 22

    Milk - 9

    Seeds and nuts - 9

**Lunch and Dinner:**

    Main Meal - 36

    Rice and Roti -  4

    Salads – 7

**Snacks:**

    Main Meal - 16

    Fruits -  22

    Sweets– 3

1. **Calorie calculation and distribution between meals**

The following steps were followed to calculate the calorie distribution between meals:

1. The following inputs are taken from the user:
   * Gender
   * Age
   * Weight
   * Height
   * Current Intensity of workout
2. Then BMR calculation is done using the following formula **[8]:**
   * BMR = 10 \* weight + 6.25 \* height - 5 \* age + 5 (for males)
   * BMR = 10 \* weight + 6.25 \* height - 5 \* age - 161 (for females)
3. Then the TDEE (Total Daily Energy Expenditure) is calculated  according to BMR and intensity of workouts:
   * Sedentary (little or no exercise) : TDEE = BMR \* 1.2
   * Lightly active (light exercise/sports 1-3 days/week) : TDEE = BMR \* 1.375
   * Moderately active (moderate exercise/sports 3-5 days/week) : TDEE = BMR \* 1.55
   * Very active (hard exercise/sports 6-7 days a week) : TDEE = BMR \* 1.725
   * Extra active (very hard exercise/sports & a physical job) : TDEE = BMR \* 1.9
4. Calorie distribution for different meals of the day:

* Breakfast - 30 %
* Lunch - 35 %
* Dinner - 35 %
* Snacks -
  + 12 % of Final calories if final calories > 2500, otherwise 10 % of final calories
  + 67 %  of the total snacks calories will be burned by workout.

1. **Steps for creating Diet Model**
2. The target calories are calculated for every meal (breakfast/lunch/snacks/dinner) using the above formula for TDEE. Percentage breakdown of macros for different users and meals [9]:
3. Protein:
4. Weight loss: 1.5 g/kg of body weight
5. Maintain/Gain:
   * + 1. Sedentary/ Lightly Active: 1.25 g/kg of body weight
       2. Rest - 1.5 g/kg of body weight

B. Carbs:

* + 1. Weight loss: 50 % of total calories
    2. Maintain/Gain:
* Sedentary/ Lightly Active: 55% of total calories.
* Rest – 60% of total calories.

C. Remaining calories will be fat.

1. The dataset is then remodified and filtered according to the user's likings and health issues selected by him/her.
2. Next, a dataset is formed consisting of cartesian products of all the datasets of a particular meal. For example, if the user has selected milk and fruits in his/her breakfast preference, then we take the cartesian product of each main breakfast item with each item in fruits and milk datasets.
3. The resulting dataset consists of different meals and their macros information. Next, every item in aparticular meal is given a learning parameter according to which the algorithm changes the quantity for that item in the given meal. For eg. Milk in breakfast is given a learning rate of 0.03. If the macros of a particular meal is not in the range of target macros, then the algorithm will increase/decrease the quantity of milk by 3%.
4. Further, a particular meal is taken from the cartesian product dataset and changes the macros of all the items in it via their learning parameter in accordance with the target macros. This process is iterated for 1000 epochs or until a meal with satisfactory macros is obtained, ie, the current meal macros must be close to the target macros.
5. Next, an error margin is considered for each macro (protein, carbs, fats) and select all the meals that are within that error margin. For example, if the error margin for all macros is 40 calories, then we find all meals whose error for all macros is less than or equal to 40 calories. Then a random meal is recommended from the final list of outputs to the user.
6. After prescribing the user with a diet plan, the user has an option to regenerate any meal from the 6 days in case he/she wants other options. This is achieved by storing the cartesian product dataframe in the database. On receiving a request to regenerate a particular meal, from the user, we recommend another random meal from the list that fits the criteria and replace it in the user’s diet plan.

**WORKOUT:**

The dataset is modified and made suitable as per needs to build a workout recommendation module which is explained further in this paper. **Each of the 88 exercises are divided into 3 categories being Run/Walk, Sports/Recreation and Gym/Exercise.**

1. **Calorie calculation and buffer**

The calories assigned for workout are calculated in the following manner:

* 33 % of calories calculated for snacks will be distributed between Breakfast, Lunch and Dinner. The remaining 67 %  of calories will be burned by workout.
* Using this calorie count, the machine algorithm will parse through the dataset and suggest workouts that meet the above criteria by burning the given number of calories.

A buffer of 20 calories is considered in consultation with Dt. Dhwani Parekh. For example, if the goal is to burn 200 calories by working out, then the model selects all those exercises whose calories burn estimation by the algorithm lies in the range 200 ± 20 i.e, 180 ≤ X (Calories burned by workout suggested) ≤ 200

1. **Steps for creating ML Model**

To develop the machine learning model for workout recommendation, the original “Calories per kg” column is firstly discarded and replaced with a new Calories per kg column named “CPK”. The CPK values for each row are calculated by first converting the values of the following 4 columns “130 lb, 155 lb, 180 lb, 205 lb” from Calories per Lb to Calories per Kg and then taking an average of these 4 columns.

Next, using the values from the above 4 columns, the coefficient “c” (which is the intercept in terms of the Linear Regression equation) is calculated for each of the 4 columns and then an  average of them is taken to get a single value of coefficient “c”.

Now using the coefficient “c” and the column “CPK” the model can predict the calories burned in one hour for a given exercise using the Equation (1):

***Y(calories burnt) = M(CPK)\*X(Weight)+C(Constant)*** (1)

Next, the **user is asked for his preferences** in regards to duration of workout he can perform on a daily basis **(The time duration considered are 30 mins, 45 mins, and 1 hour)** and the **type of workout he is interested in (Run/Walk, Sports/Recreation and Gym/Exercise)**. Next, **the above Equation (1) is applied on each row of the dataset** that satisfies the user inputs and the **Y value obtained is respectively multiplied by the time for which it is performed**. The final value of each row is then compared with the target calories required to burn **and if it lies in the buffer region** of the target calories, **the given activity is recommended to the user along with the time duration.**