Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 20 October 2024 |
| Team ID | 739652 |
| Project Title | Ai-Powered Nutrition Analyzer For Fitness Enthusiasts |
| Maximum Marks | 6 Marks |

**Preprocessing Template**

Preprocessing involves preparing data for analysis by cleaning, transforming, and organizing it. Steps include handling missing values, normalizing data, encoding categorical variables, and ensuring consistency and accuracy. It enhances the quality and usability of data for machine learning and other analytical tasks.

|  |  |
| --- | --- |
| **Section** | **Description** |
| Data Overview | Provides a summary of key nutritional data, including macronutrients, micronutrients, and dietary preferences. |
| Resizing | Adjusts input data dimensions to fit specific fitness goals, such as meal plans, portion sizes, or dietary trends. |
| Normalization | Standardizes data by converting all nutritional values into a common scale, ensuring consistent comparisons. |
| Data Augmentation | Enhances nutritional data through techniques like generating variations of meal plans using healthy substitutions. |

|  |  |  |  |
| --- | --- | --- | --- |
| Denoising | | Removes noise from nutritional data, such as irrelevant ingredients or incorrect quantities, for improved accuracy. | |
| Edge Detection | | Extracts key nutrients or combinations, like protein-carbohydrate ratios or vitamins-minerals combinations. | |
| Color Space Conversion | | Converts ingredient data into embeddings (e.g., Nutrient Embeddings) for richer analysis. | |
| Image Cropping | | Trims meal data to include only necessary components, such as primary ingredients or essential macros. | |
| Batch Normalization | | Normalizes nutritional data frequencies for balanced meal proportions, e.g., protein, carbs, fats percentages. | |
| **Data Preprocessing Code Screenshots** | | | |
| Loading Data | |  | |
| Coreleation between variables | |  | |

|  |  |
| --- | --- |
| Data Preprocessing |  |
| clean the comment \_text in both the datasets. & training and testing |  |
| Train \_test\_ split |  |

|  |  |
| --- | --- |
| Loading the pickle file |  |