Data Collection and Preprocessing Phase

|  |  |
| --- | --- |
| Date | 15 October 2024 |
| Team ID | 739652 |
| Project Title | Ai-Powered Nutrition Analyzer For Fitness Enthusiasts |
| Maximum Marks | 2 Marks |

**Data Collection Plan & Raw Data Sources Identification Template**

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

**Data Collection Plan Template**

|  |  |
| --- | --- |
| **Section** | **Description** |
| Project Overview | The objective of this project is to develop an AI-powered nutrition analyzer that helps fitness enthusiasts optimize their diets based on individual goals, preferences, and nutritional needs. The focus is on creating personalized meal plans, tracking nutritional intake, and providing real-time dietary insights. |
| Data Collection Plan | Data will be collected from various sources, including user input from fitness apps, dietary tracking platforms, nutritional databases, and food labeling APIs. The aim is to ensure comprehensive and diverse data encompassing meal plans, exercise logs, and ingredient information |
| Raw Data Sources Identified | For building the AI-powered nutrition analyzer, the primary data sources include user-generated food logs, recipes, and nutritional supplements from fitness-focused platforms like MyFitnessPal, Fitbit, and HealthifyMe. To enrich the dataset, publicly available nutritional databases (e.g., USDA Food Composition Database) and community-sourced meal plans will be integrated. |

**Raw Data Sources Template**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source Name** | **Description** | **Location/URL** | **Format** | **Size** | **Access Permissions** |
| Nutritional Data | Data on macronutrients, micronutrients, and dietary preferences | Kaggle Dataset | CSV | 2 GB | Internal Access |
| Ingredient Database | Comprehensive list of food ingredients and their nutritional values. | Repid API | JSON | 3 GB | Private (Restricted) |
| User Activity Logs | User-generated logs of meal plans and activity data | Internal Serve | CSV | 800  MB | Private |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Fitness Tracking Data | User fitness and exercise data including calories burned and workouts. | rapidapi | CSV | 100  MB | Private |