Data Collection and Preprocessing Phase

Date	21-06-2025
Team ID	SWTID1749893823
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 endeavor.

Data Collection Plan Template

Section	Description
Project Overview	The AI-powered Nutrition Analyzer helps fitness enthusiasts track nutritional intake accurately by identifying food items from images using transfer learning. It predicts nutritional content using a food database and supports real-time image-based meal logging. This enables personalized meal planning and calorie tracking to meet fitness goals.
Data Collection Plan	From Skill Wallet Guided project
Raw Data Sources Identified	The raw data sources for this project is created from the data given in the guided project in Skill Wallet (nutrition_data.csv)

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
TEST_SET	Contains labeled images similar to the training set but unseen during training. Used to evaluate the model's accuracy and	https://drive.google.com/drive/folders/1TpJpaKzYjyXIQpsEo9yDfPLrhJ4y45gr?usp=sharing	ZIP file	4.23GB	With link

	generalization on new data.				
TRAIN_SET	Contains labeled images of fruits (apples, banana, etc.) organized into folders, used to train the AI model. Helps the model learn to classify different fruit types through supervised learning.	https://drive.google.com/file/d/1aR6AwH5KtwMchyfG3PK8cth5dIPMkdu7/view?usp=sharing	ZIP file	4.23GB	With link
nutrition_data	This file contains nutritional information (like calories, protein, fat, carbs, etc.) for each fruit category. After classifying a fruit image, the model uses this data to display accurate nutritional values to the user.	https://drive.google.com/file/d/1v55iQNdpPshZ-_wUA5jbtKnCvnhbBwb1/view?usp=drivesdk	CSV	160 Bytes	Private to the session