

Data Collection and Preprocessing Phase

Date	19 th JULY 2024
Project Title	Nutrition App Using Gemini Pro

Data Collection Plan & Raw Data Sources Identification Report:

Elevate your data strategy with a comprehensive Data Collection Plan and Raw Data Sources Identification Report. This document ensures meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor for the AI Nutritionist App.

Data Collection Plan:

Section	Description
Project Overview	The AI Nutritionist App project aims to provide accurate nutritional information and analysis by leveraging AI and machine learning. The objective is to build a model that offers real-time, precise nutritional advice and calorie analysis, enhancing user health and dietary decisions. By analyzing a dataset with features such as food items, nutritional values, and user-submitted data, the project ensures informed decision-making in personal nutrition and wellness.
Data Collection Plan	<ul style="list-style-type: none">• Search for datasets related to food nutrition, calorie counts, and dietary information from reputable sources.• Prioritize datasets with diverse food items, nutritional details, and regional variations to ensure comprehensive coverage.• Utilize a mix of automated and manual data collection methods to gather accurate and up-to-date information.
Raw Data Sources Identified	The raw data sources for the AI Nutritionist App include datasets obtained from reputable food databases, public repositories, and user submissions. The provided sample data represents a subset of the collected information, encompassing variables such as food items, nutritional values

Raw Data Sources Report:

Source Name	Description	Location/URL	Format	Size	Access Permissions
UCI	Websites and blogs focused on nutrition and health providing curated nutritional data.	https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://escholarship.org/content/qt59d0d3db/qt59d0d3db.pdf&ved=2ahUKEwjrqciUsrOHAXU2SmwGHVqwBOYQFnoECCUQAQ&usg=AOvVaw2JPNuZIByoXWB rC4IW67xB	pdf	445kB	Public
UCI	Older databases previously used by the app.	https://archive.ics.uci.edu/dataset/94/spambase	CSV	14.6KB	Public