

Data Collection and Preprocessing Phase

Date	09 February 2026
Team ID	LTVIP2026TMIDS66312
Project Title	Advancing Nutrition Science through GeminiAI – NutriGen
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification

Data Collection Plan

Section	Description
Project Overview	<p>NutriGen is an AI-driven personalized nutrition assistant that generates scientifically informed diet plans and nutrition guidance based on user health inputs.</p> <p>The project does not rely on a pre-existing dataset. Instead, it dynamically processes real-time user-provided health and lifestyle information using a pre-trained Gemini AI model to generate customized nutrition recommendations.</p>
Data Collection Plan	<p>The data for NutriGen is collected directly from users at runtime through a Streamlit-based or web-based interface.</p> <p>Users provide structured health-related inputs such as:</p> <ul style="list-style-type: none"> • Age • Gender • Weight

Raw Data Sources Identified	<p>Since NutriGen uses real-time user health input and a pre-trained AI model:</p> <ul style="list-style-type: none"> • No external datasets (CSV, images, or static files) are collected. • No third-party medical databases are stored locally. • All processing occurs dynamically at runtime.
-----------------------------	--

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
User Input Data	User-entered health parameters (age, weight, conditions, goals)	Streamlit Web Interface	Structured Text / Numeric	Small	Public (user-provided)
AI Model Output	Generated personalized diet plan and nutrition guidance	Gemini Flash Lite API	Text	Variable	Restricted (API-based)