Data Collection and Preprocessing Phase

Date	05 February 2026
Team ID	LTVIP2026TMIDS66233
Project Title	Flavour Fusion: AI-Driven Recipe Blogging
Maximum Marks	6 Marks

Preprocessing Template

In the Flavour Fusion project, data preprocessing focuses on **user-provided textual input** rather than images. Since the application uses a **pre-trained generative AI model**, no traditional dataset collection or image preprocessing is required. Instead, preprocessing ensures clean, valid, and meaningful text input to generate accurate recipe blogs.

Section	**Description**
Data Overview	The data used in this project consists of user-entered text inputs such as recipe topic and desired word count. No external dataset is used.
Text Cleaning	User input is cleaned by removing unnecessary spaces and validating empty or invalid entries.
Input Validation	Ensures the recipe topic is not empty and the word count is within an acceptable range.
Token Handling	The input text is passed to the Gemini Flash Lite model, which internally handles tokenization and text processing.
Prompt Formatting	The user input is formatted into a structured prompt before being sent to the AI model for recipe generation.

Error Handling	Handles invalid inputs or API-related issues gracefully by displaying appropriate error messages.
Data Preprocessing Areas	
Loading Data	User inputs are collected directly through Streamlit text input fields.
Input Validation	Code ensures valid recipe topic and word count before processing.
Prompt Creation	Code formats the validated input into a prompt for the Gemini Flash Lite model.
Model Invocation	The formatted prompt is sent to the AI model for recipe blog generation.
Output Handling	The generated recipe text is received and displayed on the Streamlit interface.