## Optimizing E-commerce Reviews with Product-Relevant Filtering

## Overview

This project is designed to validate customer reviews by ensuring they accurately describe the correct product. Using a language model-based solution (Gemma2), the model filters reviews by accepting those that match the product description and rejecting irrelevant or mismatched content. This solution streamlines customer review analysis, improves review quality, and maintains relevance for future buyers.

## STAR Documentation

### S: Situation

E-commerce platforms often receive large volumes of customer reviews. Not all reviews, however, are relevant to the product they describe. Some mention unrelated products, use incorrect product names, or contain entirely irrelevant feedback. Inconsistent reviews can mislead potential buyers, impact purchasing decisions, and reduce the credibility of the review section.  
  
For example, on an e-commerce platform, a product page for 'Organic Almond Butter 500g' might receive a review like:

> 'Peanut butter was too salty. Not what I expected.'  
  
This project aims to create a model that will analyze reviews like this one and identify if they correctly describe the specific product page they are on. The model should accept reviews when they match the product and reject reviews that don’t, helping to improve overall review relevance and quality for shoppers.

### T: Task

The task is to develop a model that can:  
1. Reject reviews that are unrelated to the listed product.  
2. Reject reviews mentioning a different product.  
3. Accept reviews that genuinely discuss the product, even if they don’t explicitly mention the product name (e.g., valid complaints or delivery issues).  
  
By implementing this model, e-commerce platforms can automate and streamline review validation, ensuring more relevant, accurate, and helpful customer feedback.

### A: Action

The project uses a step-by-step process with the Gemma2 language model to determine if each review matches the correct product:  
  
1. Initialize the Gemma2 Model:  
 - Load the Gemma2 model with zero temperature for consistent outputs.  
 - Define a prompt template that specifies criteria for accepting or rejecting reviews based on product relevance.  
  
2. Define the Prompt Template:  
 - The prompt template includes instructions to either accept or reject a review based on its relevance to the specified product.  
 - Criteria are specified, such as accepting reviews with valid product complaints or rejecting those that are unrelated or use generic terms for different products.  
  
3. Parse Model Responses:  
 - Define response schemas for “Decision” (accept/reject) and “Reason” (brief rationale).  
 - Parse the responses to extract decision and reason for each review.  
  
4. Evaluate Reviews:  
 - For each review, the model generates a decision and reason.  
 - The model compares its prediction with manual validation data to calculate accuracy.  
  
5. Calculate Model Accuracy:  
 - The accuracy of the model is calculated against a manual validation dataset to ensure that the model maintains a high level of precision.  
  
6. Save Results and Track Performance:  
 - Results are saved into a CSV file for further analysis, including each review, the model’s decision, and the reasoning.

### R: Result

The model achieved a high accuracy rate of 95%, indicating that it successfully accepts relevant reviews and rejects incorrect or irrelevant reviews. The solution provides clear rationale for each decision, making it a useful tool for improving the quality and relevance of customer reviews.

## Examples of Model Decision-Making

Example 1:

- Product Name: Stainless Steel Water Bottle 1L

- Review: Great bottle, keeps water cold for hours!

- Decision: Accept

- Reason: Relevant product feedback.

Example 2:

- Product Name: Organic Almond Butter 500g

- Review: Peanut butter was too salty. Not what I expected.

- Decision: Reject

- Reason: Mentions a different product.

Example 3:

- Product Name: Wireless Bluetooth Headphones

- Review: The sound quality is good, but the fit is a bit uncomfortable.

- Decision: Accept

- Reason: Relevant feedback on the listed product.

Example 4:

- Product Name: Organic Green Tea Bags - 50 count

- Review: I love this tea! Refreshing and high quality.

- Decision: Accept

- Reason: Relevant product review.

Example 5:

- Product Name: Smartphone Screen Protector for Model X

- Review: Does not fit my tablet. Too small.

- Decision: Reject

- Reason: Describes a different device.

## How the Model Works

The model analyzes reviews by matching the content to the product title, rejecting unrelated or mismatched reviews. Its decision-making is guided by six main criteria, focusing on:  
  
1. Accurate product name match.  
2. Relevant product feedback (quality, usability, issues).  
3. Acceptance of minor spelling or grammar errors.  
4. Rejection of unrelated reviews, including those that mention a different product.

## Real-World Use Cases

- E-commerce Platforms: Improves the relevance of reviews for each product page. Helps streamline manual review processes, saving time and resources.

- Customer Insights and Analytics: Provides accurate data for product quality analysis, identifying genuine product-related feedback.

- Automated Moderation Systems: Ensures user-generated content adheres to platform standards, reducing noise from irrelevant reviews.

- Enhanced Shopping Experience: Increases customer trust by displaying only relevant reviews on product pages, improving the shopping experience and aiding in decision-making.