

PRODUCT REVIEW ANALYZER

1. Introduction

The Product Review Analyzer processes product reviews to perform sentiment analysis and generate responses to user queries. It uses sentiment classification (positive, negative, neutral) and a retrieval-augmented generation (RAG) approach to summarize relevant reviews.

2. Objective

The goal is to analyze reviews, provide sentiment scores, and offer summaries based on user queries about specific product features.

3. Technologies Used

- **Python, Pandas** for data processing.
- **NLTK/Regex** for text cleaning.
- **Sentence Transformers & Faiss** for embedding and retrieval.
- **GPT-2** for generating summaries.

4. Data Preprocessing

The dataset is cleaned by:

- Retaining relevant columns (Score, Summary, Text).
- Dropping missing values and duplicates.
- Cleaning review text.

5. Sentiment Analysis

Sentiment is classified using a word-based approach:

- **Positive:** Words like "good", "amazing".
- **Negative:** Words like "bad", "terrible".
- **Neutral:** When counts are equal.

6. Retrieval-Augmented Generation (RAG)

Reviews are embedded using a transformer model. For user queries, relevant reviews are retrieved, and GPT-2 generates a summary based on these reviews.

7. User Interaction

Users can query the system (e.g., "What do people say about the camera?"), and the system generates summaries based on retrieved reviews.

8. Results

The system provides accurate, context-driven summaries. For example, a query about battery life will return a generated summary based on reviews mentioning battery performance.

9. Challenges

- Handling missing values and duplicates.
- Optimizing query response times.

10. Future Work

- Improve sentiment analysis using advanced techniques.
- Optimize retrieval for better query results.
- Expand the dataset for more comprehensive analysis.

11. Conclusion

The Product Review Analyzer successfully provides valuable insights into user reviews, helping users make informed decisions based on sentiment and review summaries.