CSE508 Information Retrieval
Assignment - 4 Report: Review Summarization using GPT2

Dataset Overview
● Dataset: Amazon Food Reviews
● Total Rows: {total_rows}
● Selected Random Sample: 1200 rows

Data Cleaning Steps
1. Remove Duplicates:
● Removed duplicate rows from the dataset.
2. Handle Missing Values:
● Removed rows with missing 'Text' or 'Summary' values.
3. Text Cleaning:
● Converted text to lowercase.
● Removed URLs.
● Removed non-alphabetic characters.
● Removed extra whitespaces.
● Removed leading and trailing whitespaces.
● Remove stop words using NLTK's English stopwords list

Review Summary Generation
Model Loading
- Loaded Model Checkpoint: "results/checkpoint-500"
Summary Generation
1. Produced Summary for User Input (user_input)
2. Summary (summary)
3. Model Loading: Loaded the saved model checkpoint (checkpoint-500) using GPT2 LMHeadModel from pretrained and GPT2 tokenizer from pretrained. Set the pad token ID for the tokenizer to the end-of-sequence (EOS) token ID from the model configuration.
4. Summary Generation: Defined a function generate_summary that takes a review text as input, encodes it using the tokenizer, and generates a summary using the loaded model. The summary is generated with specific parameters such as num_beams, length_penalty, and repetition_penalty for beam search decoding.
5. User Input and Summary Generation: Prompted the user for a review text input, generated a summary using the generate_summary function, and then printed the generated summary.

ROUGE Score Calculation
ROUGE Scorer Initialization
- Metrics: "rouge1", "rouge2", "rougeL"
- Stemmer Used: True
Example Usage

- Actual Summary (provided by user): [actual_summary]
- Predicted Summary: [predicted_summary]
- ROUGE Scores: [scores]

1. ROUGE Score Calculation: Defined a function calculate_rouge_scores that takes the actual summary and the predicted summary as inputs, initializes a ROUGEScorer with ROUGE metrics ("rouge1", "rouge2", "rougeL"), and calculates the ROUGE scores using the score method of the ROUGEScorer.
2. Example Usage: Prompted the user for the actual summary, which is assumed to be provided by the user, and use the previously generated summary as the predicted summary. You then calculate the ROUGE scores using the calculate_rouge_scores function and print the scores.