Name: Aaryan S Verma
Roll number: 2020272

Report: Fine-tuning GPT-2 for Summarization on Amazon Fine Food Reviews Dataset

Introduction

In this project, we aimed to fine-tune the GPT-2 model for text summarization using the Amazon Fine Food Reviews dataset. The objective was to train a model that can generate concise summaries of product reviews. This report outlines the steps taken for data preprocessing, model training, and evaluation.

Dataset

The Amazon Fine Food Reviews dataset was used for this task. It consists of reviews of various food products along with corresponding summaries. The dataset was preprocessed to clean the text and summary columns, removing HTML tags, special characters, punctuation, and stopwords. The text was also tokenized and lemmatized for further processing.

Model Training

Tokenizer and Model Initialization: We utilized the GPT-2 tokenizer and model from the Hugging Face library.
Data Splitting: The dataset was divided into training and testing sets with a split of 75:25.
Custom Dataset Class: A custom dataset class was implemented to prepare the data for training. This class encoded the text and summary pairs and handled padding and truncation.
Hyperparameter Tuning: Hyperparameters such as learning rate, batch size, number of epochs, weight decay, and warmup steps were experimented with to optimize model performance.
Training Process: The model was trained on the review dataset using the specified hyperparameters. We employed the AdamW optimizer and utilized strategies like logging and evaluation at the end of each epoch.

Evaluation

After training, the model's performance was evaluated using ROUGE scores on the test set. ROUGE scores were computed for each predicted summary against the actual summary, measuring precision, recall, and F1-score. ROUGE scores provide insights into the model's ability to generate summaries that are close to the ground truth.

Results

The trained model demonstrated promising performance in generating summaries for Amazon Fine Food Reviews. ROUGE scores were computed to assess the quality of generated summaries compared to the actual summaries. The experiment with different hyperparameters allowed us to optimize the model's performance.

Conclusion

In conclusion, fine-tuning the GPT-2 model for summarization on the Amazon Fine Food Reviews dataset proved to be effective in generating concise and informative summaries. The model's performance can be further improved by experimenting with additional hyperparameters and training on larger datasets. Overall, this project demonstrates the potential of using transformer-based models for text summarization tasks.