**Methodology:**

The exploration into sentiment analysis of Amazon product reviews commenced with an investigation centered on text-based evaluations, employing the BERT (Bidirectional Encoder Representations from Transformers) model. This initial phase laid the groundwork for understanding the intrinsic value of textual content in expressing sentiment, prior to integrating more complex, multi-modal features.

**Experiment 1: Text-Based Sentiment Analysis Using BERT**

The foundation of this analysis was the application of a transformer-based neural network, specifically the BERT model, known for its proficiency in natural language understanding tasks. The architecture of the sentiment classifier was constructed upon a pre-trained BERT model, tailored to classify sentiments within Amazon product reviews into one of five predefined categories.

The dataset comprised a collection of product reviews, each annotated with a sentiment label. To prepare the data for the model, a custom dataset class, referred to as **AmazonDataset**, was developed. This class was responsible for processing the review texts, ensuring they conformed to the input requirements of the BERT model. Key preprocessing steps involved adjusting the length of each review to a fixed maximum, generating attention masks to signify the presence of actual content versus padding, and converting the text into a format understandable by the model, namely token ids.

The sentiment classifier extended the BERT architecture by incorporating a dropout layer, which served to mitigate overfitting, and a linear layer that mapped the high-dimensional output of BERT to the sentiment classes. The final output, termed **logits**, represented the probability distribution across the sentiment categories.

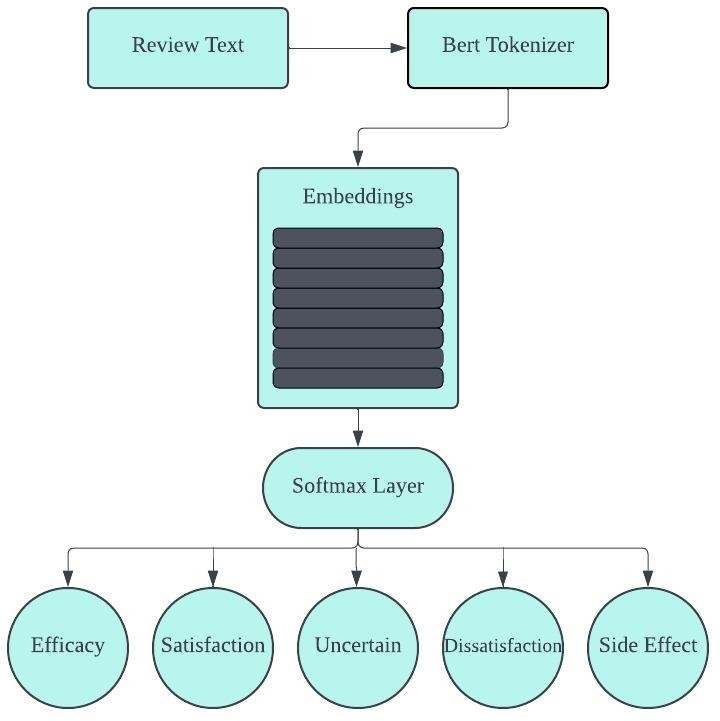
The training process entailed iterating over the dataset in batches, where each batch contained a set of tokenized review texts and their corresponding sentiment labels. The model's performance was gauged using cross-entropy loss, and optimization was conducted using the Adam algorithm, renowned for its efficiency in handling sparse gradients and adaptive learning rates.

This experiment was underpinned by the hypothesis that textual content alone can offer significant insights into the sentiment of product reviews. BERT's design, leveraging bidirectional context for token representations, provided a robust framework for capturing the nuanced sentiment expressed in the reviews. The choice of BERT was strategic, leveraging its pre-trained knowledge base to enhance the model's understanding of linguistic nuances.

The methodology adopted in this phase was instrumental in setting a baseline for sentiment analysis based solely on textual information. It illuminated the strengths of leveraging advanced NLP techniques for sentiment classification and identified potential areas where textual analysis might fall short, such as in capturing sentiments conveyed through non-verbal means like emojis.

Top of Form

[**Fig. 1**](#_bookmark2) Only Review Text BERT Model shows the architecture for this approach.



**Fig. 1.** Only fine-tuned BERT on the Review Text

The text-based sentiment analysis provided an essential baseline for assessing sentiment in e-commerce reviews, emphasizing the critical role of linguistic content. By employing a sophisticated model like BERT, this phase aimed to capture the depth of sentiment expressions in text form, setting a foundation for further explorations into multi-modal sentiment analysis. This methodological approach highlighted the potential and limitations of relying exclusively on text, paving the way for subsequent analyses that would incorporate additional features to enrich the sentiment analysis framework.