5.1. The dataset used for sentiment analysis consists of product reviews. Each review includes text describing the reviewer's opinion about a product. The dataset may contain various attributes such as product ratings, review titles, and timestamps. For this analysis, we focused on the 'review.text' column, which contains the text of the reviews.

5.2. The processing steps involved in sentiment analysis include:

- Loading the dataset: We loaded the dataset into a Pandas DataFrame, ensuring that it contains the necessary columns for analysis.
- Data Cleaning: We cleaned the text data by removing punctuation, converting text to lowercase, and removing stop words using the spaCy library.
- Sentiment Analysis: We implemented a sentiment analysis model using spaCy to analyse and classify the sentiment of each product review. The model calculates the average sentiment score of the text and classifies it as positive, negative, or neutral based on the score.

5.3. The sentiment analysis results were evaluated based on a sample of product reviews. Each review was classified as positive, negative, or neutral. The accuracy of the sentiment analysis model was assessed by comparing the predicted sentiment with the reviewer's opinion. The evaluation showed that the model was able to accurately predict the sentiment of most reviews. However, there may be cases where the model misclassifies the sentiment due to the complexity of natural language and subjective interpretations.

5.4. Strengths:

- The model efficiently processes large volumes of text data, making it suitable for analysing large datasets of product reviews.
- The model can be customized by adjusting parameters and fine-tuning the text cleaning and sentiment analysis process to improve accuracy.
- The model can be scaled to handle different types of text data and sentiment analysis tasks, making it adaptable to various use cases.

Limitations:

- Sentiment analysis is inherently subjective and may vary based on individual interpretations of text.
- The model may not accurately capture the sentiment of all reviews, especially those with ambiguous language.
- The accuracy of the sentiment analysis model depends on the quality and relevance of the dataset. Low-quality or biased data may result in inaccurate predictions.
- The model may be limited to analysing text in a specific language or may not perform well with text containing slang, abbreviations, or informal language.