2.3 Data Annotation Details

Customer reviews from the Lazada website were gathered, with a focus on mask products, and data annotation was done to categorize them.

2.3.1 Labels, Classes and Annotations Guidelines

Reviews with Negative Sentiments:

Those that voice discontent or critique the mask items, pointing up things like inadequate quality or inefficiency. For example, using words or phrases like "bad," "low quality," "no good," and other negative descriptions, are considered negative sentiments. The star count for negative review would be 1 and 2.

Reviews with Neutral Sentiments:

Reviews that offer unbiased commentary, noting both advantages and disadvantages without favoring one over the other. The star count for neutral review would be 3.

Reviews with Positive Sentiments:

Reviews that stress positive features of the mask items, like its effectiveness, quality, or value, and convey satisfaction or appreciation for them. For example, words or phrases like "nice", "good", "comfortable", "fulfill" and other positive keywords. The star count for positive review would be 4 and 5.

2.3.2 Target Variable

Target Variable Name: Sentiment

Description: The target variable represents the sentiment classification of the review, used to train the model to predict sentiment based on the review text.

Classes and Values:

- **0**: Negative Sentiment (indicating dissatisfaction or criticism)

- **1**: Neutral Sentiment (indicating balanced or impartial feedback)

- **2**: Positive Sentiment (indicating satisfaction or praise)

2.3.3 Filtering and Final Data

- The dataset is filtered to exclude records with a neutral sentiment (Sentiment = 1).