Sentiment Analysis Report

Dataset Description:

This dataset contains product reviews from Amazon. It consists of various products across different categories and includes textual reviews provided by users along with corresponding ratings.

Details of Preprocessing Steps:

The preprocessing steps include tokenization, stop word removal, and lemmatization. Each review text is normalized to lowercase, tokenized into words, and then stop words are removed. Finally, the remaining words are lemmatized to their base form.

Evaluation of Results:

The sentiment analysis model was tested on a sample of product reviews. However, all sentiments were classified as 'Neutral', indicating that the model may not be accurately capturing the polarity of the reviews. Further investigation is required to improve model performance.

Insights into Model's Strengths and Limitations:

The model's strength lies in its simplicity and ease of use. It leverages SpaCy's pre-trained model for natural language processing tasks, making it suitable for quick sentiment analysis tasks. However, the model's limitation is evident in its inability to accurately classify sentiments in the given dataset. This may be due to the dataset's characteristics or the model's lack of fine-tuning specifically for sentiment analysis tasks.