

Sentiment Analysis Report

Dataset Description

The dataset used is the Amazon Consumer Reviews dataset from Datafiniti which includes customer product reviews. It contains the customer opinions that were analyzed for sentiment.

Preprocessing Steps

To prepare the text for analysis, several cleaning steps were applied:

- Text was changed to lowercase.
- Extra spaces were removed.
- Common words like “*the*”, “*is*”, and “*of*” were removed.
- Only alphabetic words were kept, removing numbers and symbols.

Evaluation of Results

The sentiment analysis model classified each review as either positive, neutral or negative. The polarity score ranged from -1, being negative to +1, being positive.

Most reviews were positive, showing that customers are generally satisfied with their purchases. A smaller number of reviews were Negative or Neutral, suggesting some mixed reviews.

Insights and Limitations

Strengths:

- The model quickly sorts reviews into clear categories.
- Results can help businesses understand customer satisfaction trends.
- Preprocessing reduces noise from irrelevant text.

Limitations:

- The model struggles with sarcasm or mixed opinions, lacking context.
- The use of polarity scores can oversimplify complex emotions.

This report was written by : Jessica Ramathibe