

# Sentiment Analysis of Amazon Product Reviews

## Introduction

As a data analyst student, I embarked on a capstone project to explore the realm of sentiment analysis, focusing on Amazon product reviews. This project aimed to apply data analytics techniques to extract insights from customer feedback, enhancing my understanding of natural language processing and sentiment analysis.

## 5.1 Dataset Description

The dataset consists of product reviews collected from Amazon. It includes various fields such as review text, rating, and other metadata. For this analysis, the primary focus was on the "reviews.text" column, which contains the text of the customer reviews. This dataset provides a broad spectrum of customer feedback on products, making it suitable for sentiment analysis.

## 5.2 Preprocessing Steps

Preprocessing involved several steps to clean and prepare the text data for analysis:

- Removal of stop words and punctuation to reduce noise and focus on meaningful words.
- Lowercasing all text to ensure consistency.
- Tokenization, converting text into manageable pieces or tokens for analysis.
- These steps helped in refining the dataset for more accurate sentiment analysis.

## 5.3 Evaluation of Results

The sentiment analysis revealed an average polarity score of approximately 0.38, indicating a generally positive sentiment across the reviews. The average subjectivity score was around 0.57, suggesting a mix of subjective and objective reviews. The distribution of polarity scores was skewed towards positive values, highlighting the prevalence of positive feedback in the dataset.

## 5.4 Model's Strengths and Limitations

**Strengths:** The model effectively identifies general sentiment trends within the reviews, providing valuable insights into customer opinions and product perception.

**Limitations:** The model may not capture the nuances of sentiment in more complex reviews, such as those containing sarcasm or mixed emotions. Additionally, the analysis is limited by the quality of preprocessing and the inherent biases in the dataset.

## Sentiment Distribution

**Average Polarity:** The average polarity score is approximately 0.38, indicating a generally positive sentiment across the reviews. This suggests that the majority of customers have favorable opinions about the products reviewed.

**Average Subjectivity:** The average subjectivity score is around 0.57, suggesting that reviews tend to be somewhat subjective but not overly so. This implies that while personal

opinions and experiences influence the reviews, they also contain objective information about the products.

## **Visualization**

The distribution plots provide further insights into sentiment across the reviews:

- **Polarity Distribution:** The histogram of polarity scores shows a skew towards positive values, confirming that positive sentiments predominate in the product reviews. There's a peak in the positive range, indicating a concentration of reviews with highly positive sentiments.
- **Subjectivity Distribution:** The histogram of subjectivity scores shows a broad distribution, indicating a mix of subjective and objective reviews. The peak in the mid-range suggests that many reviews combine personal opinions with factual descriptions of the products.

## **Recommendations**

Based on these findings, the following recommendations can be made:

- **Leverage Positive Feedback:** The overall positive sentiment suggests that customers are generally satisfied with the products. Marketing strategies could highlight the features and aspects most often associated with positive reviews to attract new customers.
- **Address Common Criticisms:** Although the sentiment is generally positive, analyzing the content of more negative reviews can help identify areas for product improvement or customer support enhancements.
- **Encourage Balanced Reviews:** Since the reviews show a mix of subjectivity and objectivity, encouraging users to provide balanced reviews containing both personal opinions and factual descriptions can help prospective customers make informed decisions.

## **Conclusion**

The sentiment analysis of product reviews reveals a predominantly positive customer sentiment and provides valuable insights into how customers perceive the products. By focusing on the strengths highlighted in positive reviews and addressing the issues raised in negative ones, product developers and marketers can further improve customer satisfaction and product appeal.