

Sentiment Analysis for Brand Monitoring

Project Overview

This project analyses customer sentiment based on Amazon product reviews to understand brand perception, extract key themes, and provide actionable insights for improving customer satisfaction.

Business Problem

Understanding customer sentiment is critical for brand management and product improvement. Amazon Fine Foods receives thousands of customer reviews, but this data is underutilized without proper analysis. This project aims to leverage Natural Language Processing (NLP) to analyze customer feedback and extract actionable insights.

Objectives

- Assess customer sentiment to understand brand perception.
- Identify key aspects of products that drive positive or negative sentiment.
- Track changes in customer sentiment over time.
- Provide strategic recommendations to improve product offerings and marketing efforts.

Dataset

The dataset used in this analysis is the **Amazon Fine Food Reviews** dataset, which contains over 500,000 customer reviews of fine food products. A sample of 100,000 reviews was used for this project for efficiency.

Features

- **UserId**: Unique identifier for each user.
- **ProductId**: Unique identifier for each product.
- **Score**: Product rating (1-5).
- **Text**: Full review text.
- **Time**: Timestamp of the review.

Methodology

This project follows a structured approach, involving:

1. Data Collection and Cleaning

- **Data Cleaning:** Handled missing values and removed duplicates to ensure data quality.
- **Text Preprocessing:** Applied text cleaning, tokenization, stopword removal, and lemmatization to prepare the data for analysis.

2. Exploratory Data Analysis (EDA)

- Conducted descriptive analysis and visualized the **rating distribution** and **review length**.

3. Sentiment Analysis

- **VADER Sentiment Analyzer** was used to classify reviews as **Positive**, **Negative**, or **Neutral**.

4. Topic Modeling

- Applied **LDA (Latent Dirichlet Allocation)** to identify the key themes in reviews, helping understand customers' main areas of focus.

5. Time Series Analysis

- Analyzed **sentiment trends over time** to identify any seasonal variations in customer sentiment.

Results

Sentiment Analysis Summary

- **Positive Reviews:** 91,301
- **Negative Reviews:** 6,722
- **Neutral Reviews:** 1,805

Key Insights

- Sentiment distribution shows a high prevalence of **positive sentiment**, indicating overall satisfaction.
- **Topic Modeling** revealed key themes such as **taste**, **quality**, and **price**.
- **Time Series Analysis** shows fluctuations in customer sentiment, suggesting seasonal factors or product changes affecting satisfaction.

Conclusions and Recommendations

Conclusions

- The majority of customer reviews are positive, indicating a high level of satisfaction.
- Negative reviews highlight areas needing improvement, particularly related to specific features and quality.

Recommendations

- **Product Quality:** Address recurring issues from negative reviews to enhance product quality.
- **Marketing Campaigns:** Highlight positive features such as taste and quality in advertisements.
- **Customer Engagement:** Set up proactive sentiment monitoring to track customer satisfaction and address emerging issues quickly.