PROJECT PROPOSAL

Starbucks Sentiment Analysis on Aspect-Based Customer Reviews

Members

Deepu Gondhi	Kaushik Apoori	Manumitha Kamarapu	Teena Poorna Sudheera Maganti
11737274	**11705023**	**11684243**	**11755850**

Github link: https://github.com/kaushikapoori/Group_4_NLP

1.Motivation

The project aims to utilize sentiment analysis on Starbucks customer reviews to identify key areas for improvement. The huge volume of reviews data collected through customer feedback is difficult for the company to extract actionable insights efficiently. The traditional methods such as manually analyzing feedback are time-consuming and may risk missing critical patterns or trends. To overcome this problem, automating the analysis of customer sentiments, can provide actionable insights that can enhance customer service, optimize operational costs, and solve concerns raised. In this way, the proposed approach will help in identifying strengths and weaknesses in various areas of the business, such as customer service, product quality, store environment, and pricing strategies. By such analysis, Starbucks can make data-driven decisions and helps them improve overall customer experience and operational efficiency.

2. Significance

By identifying the areas of dissatisfaction or by understanding the customer's sentiment about service quality and operation, it benefits the Starbucks to make targeted improvements. Identifying patterns in customer feedback such as, understanding product preferences will make them bring changes to the menu and products, aligning with the customer interests. Furthermore, feedback related to the store environment can help them optimize operational efficiency, benefiting Starbucks in reducing costs and improving store performance. In this way, sentiment analysis on customer reviews not only solves day-to-day concerns but also supports bringing strategic decisions. Ultimately, the sentiment analysis done will enable Starbucks to make data-driven decisions that improve customer satisfaction, enhance operational improvements, and maintain a competitive edge in the market.

3. Objectives

The main goals of this project are:

- To identify key sentiments and trends among various aspects such as service quality, product offerings, store ambience, and pricing.
- To utilize NLP techniques and categorize sentiments across a large volume of reviews.
- To create a machine learning model and classify customer reviews into positive, negative, and neutral sentiments.
- To produce detailed visualizations that summarize key findings, trends, and actionable recommendations, providing a clear roadmap for implementing improvements.
- The success criteria for the project includes achieving a sentiment classification accuracy of at least 80%.
- To extract at last five actionable insights across different aspects of the business.
- To Deliver clear, actionable reports with practical recommendations for business improvements.

4. Features

In this project, we are mainly focusing to

- Technical Aspects: Identify sentiments around specific aspects of the Starbucks experience, such as service quality, product taste, store ambiance, and pricing incorporating Aspect Based Sentiment Analysis.
- Utilize NLP techniques and extract common keywords, topics, and trends from customer reviews, which ultimately will help us in highlighting frequently mentioned issues or praised areas.
- Perform sentiment analysis and map them geographically to visualize customer feedback from different store locations, identifying regional trends and location-specific issues.

Deliverables and Technical Milestones:

- We will deliver a trained sentiment analysis model for classifying review sentiments
- Will deliver reports and visualizations that highlight key themes and provide actionable insights.
- One of the technical milestones we could face is to handle missing or incomplete reviews and establishing strategies for excluding such data.
- Another milestone we could face is to, detect and remove duplicate reviews.

5. Dataset

Material for this project is based on the Starbucks Reviews Dataset, which was web-scraped from the Consumer Affairs website-a comprehensive collection of consumer reviews and ratings for Starbucks.

Size: The dataset contains the reviews_data.csv file, which is 464.6 kB in size, and contains 741 reviews in total. This would form a good basis for extracting aspects and their sentiments.

Attributes from the dataset:

- Name: Reviewer's name.
- Location - The location or city associated with the reviewer, if provided.
- Date: The date when the review was posted.
- Rating: The rating given by the reviewer out of five stars.
- Review: The textual content of the review that depicts the experience and opinions of the reviewer.
- Image Links: Attach links to images with the review, if any.

Preprocessing:

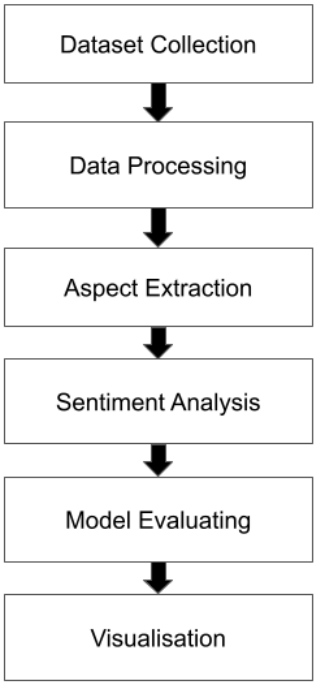
Handling of Missing Values: In cases of missing or incomplete entry in name, location, date, ratings, and review, values should handle the missing value by either removing or imputing it accordingly.

Text Preprocessing: removing stop words, punctuation, special characters, and lemmatizing into base forms.

Timestamp Transformation: Normalizing the dates to allow for temporal trend analysis.

6. Visualization

The overall process of the project is as follows:



This is how the result looks like. The data shown below are example

Aspect Extraction Result

Review ID	**Category**	**Key Insight**
401	Coffee	Coffee is amazing
402	Service	Service is bad
403	Atmosphere	Ambiance is cool

Sentiment Analysis Result

Aspect	**Positive**	**Negative**	**Neutral**
Coffee	85%	10%	5%
Service	75%	15%	10%
Atmosphere	80%	10%	10%

We also provide graphs of the results for better understanding.