Internship Task

Intern Infotech Virtual Learning Internship Program

Intern Infotech

Problem Statement:
Understanding Customer Sentiment and Insights through Starbucks Review Analysis

Data Description:
- Name: The reviewer's name, if available.
- Location: The location or city associated with the reviewer, if provided.
- Date: The date when the review was posted.
- Rating: The star rating given by the reviewer, ranging from 1 to 5.
- Review: The textual content of the review, capturing the reviewer's experience and opinions.
- Image Links: Links to images associated with the reviews, if available.

Background:
This dataset comprises a comprehensive collection of consumer reviews and ratings for Starbucks, a renowned coffeehouse chain. Acquired through web scraping from the Consumer Affairs website, it offers valuable insights into customer sentiment and feedback regarding various Starbucks locations. The dataset presents a rich source of data for conducting sentiment analysis, extracting consumer insights, and leveraging natural language processing (NLP) techniques for diverse tasks.

Objective:
The objective of this internship project is to analyze Starbucks reviews using natural language processing (NLP) techniques to extract customer sentiment, identify consumer preferences, and derive actionable insights for Starbucks locations.

Key Components:
1. Data Exploration: Familiarize with the dataset, examining the distribution of ratings, exploring textual reviews, and identifying any patterns or trends present.

2. Preprocessing: Cleanse and preprocess the textual reviews, handling punctuation, stopwords, and other noise, while retaining essential information for analysis.

3. Sentiment Analysis: Utilize NLP techniques to perform sentiment analysis on the reviews, categorizing them as positive, negative, or neutral based on the sentiment conveyed by the text.

4. Consumer Insights: Extract insights from the reviews regarding customer preferences, common complaints, areas for improvement, and notable positive experiences. Explore correlations between ratings, sentiments, and reviewer demographics (e.g., location).

5. Topic Modeling: Employ techniques like Latent Dirichlet Allocation (LDA) or Non-Negative Matrix Factorization (NMF) to identify prevalent topics within the reviews, uncovering themes and discussions relevant to Starbucks customers.

6. Visualization and Reporting: Create informative visualizations, such as word clouds, sentiment distribution plots, and topic clusters, to present the findings. Generate comprehensive reports summarizing the analysis and actionable insights for Starbucks management.

Expected Outcomes:
- Insights into customer sentiment towards Starbucks locations, including common themes and sentiments expressed in reviews.
- Identification of key areas for improvement and opportunities to enhance customer satisfaction and experience.
- Visualization of review sentiments and topics to facilitate easy interpretation and communication of findings.

Deliverables:
- Preprocessed dataset.
- Sentiment analysis results.
- Consumer insights and actionable recommendations report.
- Visualizations summarizing review sentiments and topics.

Conclusion:
This project aims to leverage natural language processing (NLP) techniques to analyze Starbucks reviews, providing valuable insights into customer sentiment and preferences. By understanding customer feedback, Starbucks can make informed decisions to enhance customer satisfaction, drive loyalty, and improve overall business performance.