Analysis of TripAdvisor Dataset on Restaurant Reviews

The Analysis focuses on reviews given for restaurants in Barcelona collected from TripAdvisor. It is conducted by adapting Natural Language Processing Techniques to adapt a meaningful interpretation by analyzing the sentiment of the reviewers on the restaurants.

CONTEXT:

Online Reviews have a huge impact on Restaurant and Cafe Businesses. Platforms like TripAdvisor play a crucial role to impact a Customer's Decision and a Restaurant's Reputation. The Reviews posted online offer both opportunities and risk, so it is important to analyze the Sentiment of the reviewers. This might contribute to making Restaurateur's business decisions more precise.

——

CRITERIA FOR SUCCESS:

Our goal is to find out the Sentiment behind the reviews on the Restaurants in Barcelona city posted by the reviewers. This is a form implementing Supervised Learning (Sentiment Analysis) based on Keywords (Topics) derived from Unsupervised Learning (Topic Modelling).

——

SCOPE OF SOLUTION SPACE:

- How can the sentiment of the reviewers be categorized based on Topics/ Keywords (Topic Modelling?
- Do Customers' Demography contribute to the sentiment of the reviews? (e.g. Location, Family Size, etc)

——

CONSTRAINTS WITHIN SOLUTION SPACE:

- No Clear Data Dictionary Available.
- Needs faster and larger memory capacity to run the analysis.
- Data Preprocessing and Inference is difficult.

——

STAKEHOLDERS TO PROVIDE KEY INSIGHTS:

- BI team of corresponding Restaurateurs in the cities of Barcelona, London, Madrid, New Delhi, New York and Paris
- Customer Insights Dept, TripAdvisor

——

KEY DATA SOURCES:

- Web Scraped Data from Online Reviews Posted on TripAdvisor.
- Dataset posted on Kaggle.
- Dataset sourced and cited from [2205.01759] Explain and Conquer: Personalised Text-based Reviews to Achieve Transparency (arxiv.org)

——