Data 620 - Final Project Proposal
Bikram Barua & Leticia Salazar
May 9, 2023

Proposal: Sentiment Analysis on Trip Advisor Hotel Reviews

Motivation:

- Everyone loves travelling and hotels play a crucial role in traveling. To plan your trip better, you want to find the best choice given your preferences.
- You can explore reviews to find a great place to rest and use this to plan your travel journey.

Data Source:

The data will be obtained from **Kaggle** that consists on 20k reviews from Trip Advisor.

- Dataset has two columns Review and Rating.
    - Review column has the text of each review
    - Rating column has rating values in integer form ranging from 1 to 5 (stars)

Project Goals:

1. Identify the links between key travel related phrases and their related ranking and sentiment.
2. Evaluate the relationship between the text contents and the sentiment rating.

Statistical Analysis:

1. Word frequency count
2. Percentage rating
3. Word cloud by rating
4. Review length
5. Top 20 most common words

Sentiment Analysis:

Using sentiment analysis, we will attempt to classify reviews as negative, neutral, or positive. We will use natural language processing to examine the following questions:

1. Do useful review attributes affect rating?
2. Does review text affect the review attribute?
3. Can review text determine if a hotel receives a high or low rating (stars)?

Work plan:

- Proposal → Bikram and Leticia
- Data collection / load → Bikram and Leticia
- Data pre-processing → Bikram and Leticia
- Data cleaning → Bikram and Leticia
- Statistical analysis → Bikram and Leticia
- Data Visualization → Bikram and Leticia
- Building a Machine Learning Model / classification → Bikram and Leticia
- Sentiment Analysis → Bikram and Leticia
- Conclusion → Bikram and Leticia
- Review → Bikram and Leticia
- Presentation → Leticia

Concerns:

- Simultaneously working on GitHub
- Managing workflow in a timely manner
- Unexpected malfunction of computer / apps

References:

- Data Source LARXEL. (2020, September 30). *Trip advisor hotel reviews*. Kaggle. Retrieved April 27, 2023, from https://www.kaggle.com/datasets/andrewmvd/trip-advisor-hotel-reviews?resource=download