Social Media Analysis Project Proposal

Group Composition:

1. Samheeta Gourammolla
2. Rahul Ishwar Pulluri
3. Aditya Sambhaji Jadhav

Dataset: Rotten Tomatoes Reviews for Online Streaming Shows
Dataset link: Rotten Tomatoes Reviews for Online Streaming Shows (kaggle.com)

Dataset Summary:

The dataset contains information about movies, including various attributes such as the percentage of critics who gave positive reviews, the percentage of audience members who gave positive reviews, the title of the movie, the rating on a scale from 0 to 5, the review text, the sentiment of the critic (1 for positive, 0 for negative), and possibly additional features.

- Percentage of Critics Positive Reviews: The percentage of critics who provided positive reviews for the movie.
- Percentage of Audience Positive Reviews: The percentage of audience members who provided positive reviews for the movie.
- Title: The title or name of the movie.
- Rating: The numerical rating of the movie on a scale from 0 to 5.
- Review Text: The text of the review provided by either critics or audience members.
- Sentiment: The sentiment label assigned to the review text, indicating whether it is positive (1) or negative (0).

Insights into the network:

The "network" column in our dataset can be considered as a network, specifically a network of streaming platforms and the movies available on each platform. In this network, the nodes represent the streaming platforms (e.g., Netflix, Hulu, Paramount+), and the edges represent the movies available on each platform.

- Nodes: Streaming Platforms: Each streaming platform (e.g., Netflix, Hulu, Paramount+) is represented as a node in the network.

- Edges: Platform-Movie Relationships: Edges connect each streaming platform node to the movies available on that platform. For example, there would be an edge from the Netflix node to each movie available on Netflix, indicating the presence of that movie on the Netflix platform. Similarly, there would be edges from other platform nodes (e.g., Hulu, Paramount+) to the respective movies available on those platforms.

Problem:

This project focuses on developing a robust sentiment analysis and recommendation system tailored for movies, drawing insights from critic and audience reviews. Leveraging advanced natural language processing techniques, the sentiment analysis component classifies movie reviews into positive or negative sentiments, facilitating a deeper understanding of audience perceptions.