Sentiment Analysis of Movie Reviews

As a movie enthusiast I often find myself spending hours together scrolling through different tiles from various streaming platforms to find the movie that might be interesting for me. Although search engines like google and Bing, as well as streaming platforms like Netflix, Amazon etc. do provide some recommendations based on the movies I search for or watch on the platform, I wanted to understand how exactly they recommend. With this project I aim to analyze the sentiment of what reviewers are saying about a movie and make a recommendation of whether someone should watch a movie.

Data

Movie Similarity data: [Sentiment Analysis of movie review | Kaggle](#)

Approach

1. Data Cleaning
   a. Identifying the data types of each feature
   b. Updating the data types if wrongly identified while loading the data
   c. Identify rows and columns with missing values
   d. Univariate and Bivariate analysis on columns
2. Feature Engineering
   a. Fix Missing values
   b. Eliminating Columns with identity information
   c. Grouping Column values based on certain orders or categories that make more sense
3. Model Building
   a. Implement a Logistic Regression for Sentiment classification of Movie Reviews
   b. Based on the Sentiment classification verify if IMDB rating is accurate to sentiment classification.
4. Writing an Article explaining the methodology followed in making the predictions.