**Sentiment Analysis of Fan Forums in the Entertainment Sector**

Pooja Joshi & E.Code [E25007]

# Overview :

# The objective of this report is to detail the process and findings from the sentiment analysis conducted on fan forums within the entertainment sector. The analysis aimed to identify key themes and sentiments expressed by viewers about various films, utilizing techniques such as text preprocessing, topic modeling, and visualization.

# Objective:

The primary goals included:

* Preprocessing fan forum reviews to clean and prepare data for analysis.
* Conducting sentiment analysis using LDA to identify prevalent themes in reviews.
* Visualizing findings through word clouds and sentiment distribution charts.

# Assigned Task(s) :

· Sentiment analysis of fan forums.

· Topic modeling using LDA.

· Visualization of results through word clouds and bar charts.

# Task Details :

**Task 1:** Sentiment Analysis of Fan Forums

* **Status:** Completed
* **Details:**
  + **Data Cleaning:**
    - Removed stop words, punctuation, and irrelevant characters from the reviews using NLTK.
    - Handled missing values by filling them with appropriate placeholders.
  + **Sentiment Analysis:**
    - Applied LDA (Latent Dirichlet Allocation) to identify 5 main topics from the fan forum reviews.
    - Analyzed the topics to interpret viewer sentiments, categorizing them as positive, negative, or mixed.
  + **Visualization:**
    - **Word Cloud:** Created a word cloud to visualize the most frequently used words in the reviews.
    - **Sentiment Distribution Bar Chart:** Plotted the average sentiment scores for each identified topic using a bar chart to represent the sentiment distribution visually.
    - **Topic Analysis Table:** Compiled a table summarizing the keywords from each topic and their corresponding average sentiment scores for easier interpretatio

# 

# 

# 

# 

# 

# 

# 

**Progress :**

· **Accomplishments:**

* Successfully cleaned and preprocessed the data.
* Identified key topics from the reviews using LDA.
* Generated visualizations (word cloud and sentiment distribution) to represent findings.

· **Metrics:**

* Topics identified: 5
* Number of reviews analyzed: [Insert total number of reviews]
* Average compound sentiment score for each topic: [Insert average scores]

# Challenges and Solutions :

· **Challenges Faced:**

* Difficulty in preprocessing due to varied text formats and language usage in reviews.
* Ensuring accurate interpretation of topics from LDA results.

· **Solutions Implemented:**

* Utilized NLTK for efficient text preprocessing, including stopword removal and punctuation handling.
* Engaged in thorough analysis of LDA output to ensure proper understanding of topics.

# Next Steps :

· **Upcoming Tasks:**

* Conduct further analysis on audience metrics and correlate findings with sentiment results.
* Prepare a presentation summarizing key insights and recommendations based on the analysis.

· **Goals:**

* Aim to finalize the analysis report by [insert date].
* Set targets for enhancing the accuracy of sentiment analysis using advanced techniques, if applicable.

# Conclusion :

* **Summary**: This report outlined the process of conducting sentiment analysis on fan forums in the entertainment sector. The key steps included data preprocessing, topic modeling, and visualization of sentiment findings. The identified topics reflect a mix of sentiments, providing valuable insights into audience perceptions.

# · **Acknowledgments**: Thank the audience for their time and attention.

# Instructions:

1. Use Google Docs. Single Column
2. TNR stands for Times New Roman: B - Bold
3. Use images as required with proper references
4. Use charts, tables as per your requirement.
5. Number of Pages: 2 to 8 for each task report.