



Movie Recommendations

Aashi, Akshaya, Harshit, Alisha, Tyler

Table of Contents



01

Problem Statement

02

Pre - Processing

Data sources, Frequency
of Important Words

03

Analysis

Lift Scores, Topic Modeling,
Sentiment Analysis

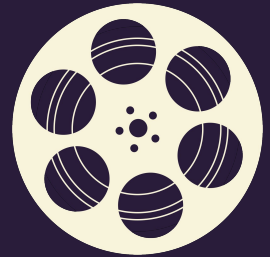
04

**Insights and
Recommender System**

User generated feedback

01

Problem Statement



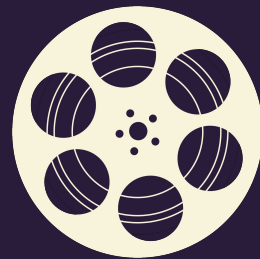
Problem Statement

- Thousands of movies across multiple streaming platforms... How do you know which one to pick to watch?
- Every movie watcher has a different opinion
- Create a recommender system by scraping IMBD.com to filter movie reviews and ratings
 - Generates recommendations based on user input attributes



02

Pre-Processing



Data Pre-Processing

- IMDB.com
 - Scraped 10,000 movies and TV shows along with ratings
 - Movies/shows rated on a 1-10 scale



Data Pre-Processing

- Majority of descriptive reviews feature action, horror, drama, and comedy
 - Sharp drop off for word frequency for the rest of the genres
- Can conclude that most watched movies reside in those four genres

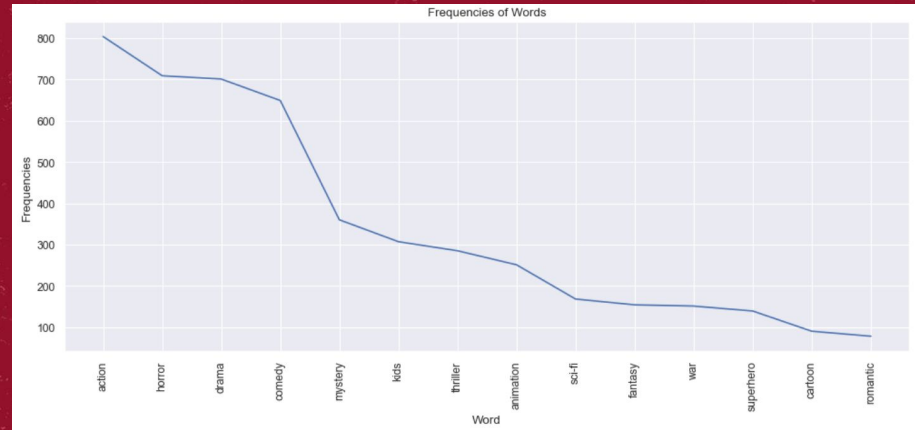
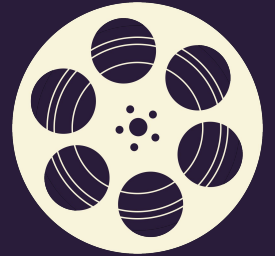


Figure 1 : Word Frequency Count for Genres

03

Data Analysis



Analysis

- Animation has highest lift scores with Romantic, Superhero, Kids, & Cartoon
- Romantic has highest lift scores with Drama, Animation & Comedy
 - Zero lift scores with cartoon and thriller genres
- Sci-fi has highest lift scores with Mystery & Fantasy
- Overall, lift scores are accurate to prior assumptions about movie genres

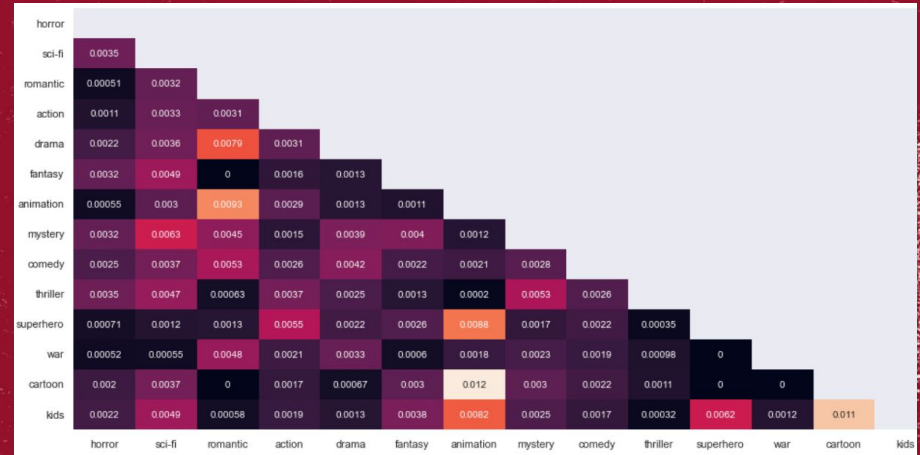


Figure 2 : Lift values between different genres

Analysis

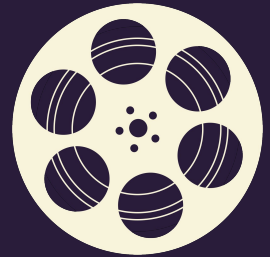
- Mystery, Comedy, Action, Sci-fi, and Drama are clustered
 - These genres often overlap within movies
 - Inception
- Animation, Kids, Superhero, and Sci-fi are grouped closely together
 - Kids movies are often animated



Figure 3: MDS Of How Different Genres Are Related

04

Insights and Recommender System



Insights and Recommender System

- Takes list of attributes that the user wants to see in a movie
- Runs through the table of movies and their reviews & ratings
 - Table was generated from scraping IMDB website
- Returns the top 5 movies and their cosine similarity for what the user is looking for based on their input



Insights and Recommender System

- User wants a movie with the following attributes:
 - Scary, Funny, Horror
- Recommender System generates top 5 movies that user would be subject to watch and enjoy
- **Business Benefit:** To help streaming platforms generalize their content more efficiently for the customer

	Movie	Cosine_Similarity
209	trick 'r treat	0.065360
110	my best friend's exorcism	0.061109
18	barbarian	0.052833
148	sissy	0.052495
151	speak no evil	0.051136

Thank You!



Questions?