

MY MOVIES

NATI MARCUS

OUTLINE

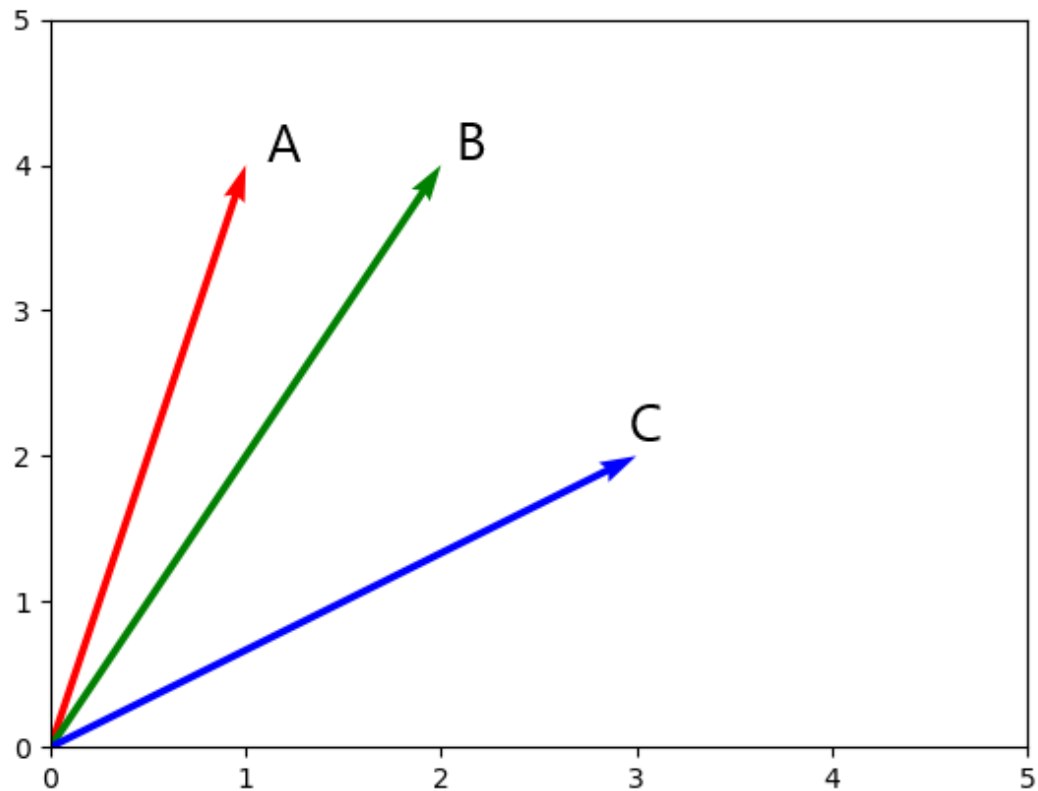
- Problem Statement
- Background
- Data
- Methods & Modeling
- Results & Conclusions
- Future Research
- App

PROBLEM STATEMENT

All major TV and movie streaming platforms today have their own recommendations built into their software. However, each streaming service is only able to recommend whichever movies are currently on its respective platform. To solve this problem, I will build a recommender system model for movies on Netflix, Hulu, Amazon Prime Video, and Disney+ to predict which movies to recommend to a user in one location. I will use movie description, actors, genre, and director data to compare movies. The model will be evaluated based on the cosine similarity scores between titles in the dataset.

BACKGROUND

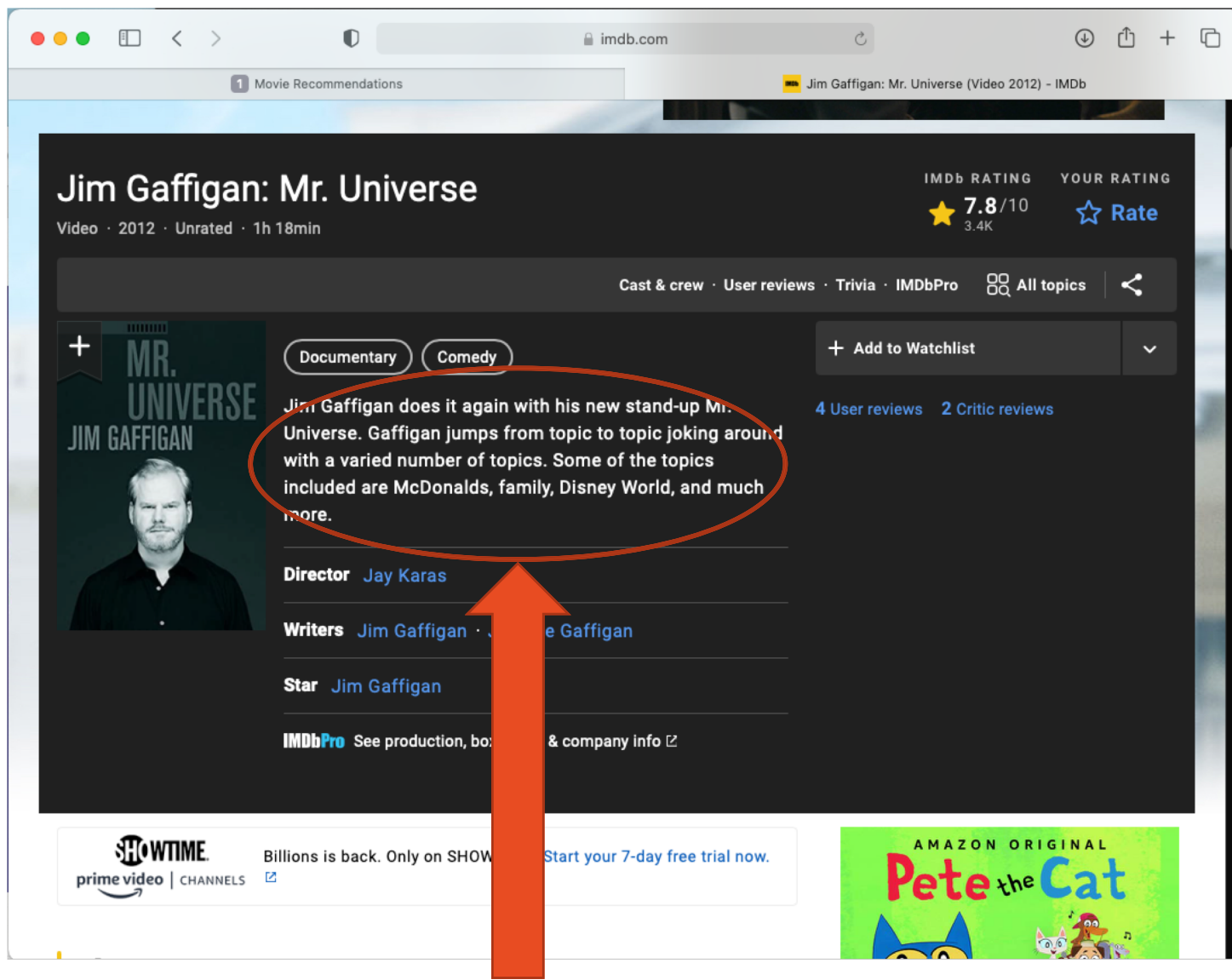
- Recommender systems are machine learning models that predict content to recommend to a user, whether it be for movies, online purchases, or other relevant content.
- Cosine similarity is often used to determine how similar one movie is to another or how similar the users are based on the movies they consume
 - I will be comparing movie to movie in this analysis



DATASET

- The original dataset came from Kaggle
- Contained every title from Netflix, Hulu, Prime Video, and Disney+
- Features included, title, year made, age range of movie, director, IMDB score, rotten tomatoes critic score, and which platform the movie was on






DATA COLLECTION

- Created links for each movie to webscrape
- Webscraped movie description from IMDB link

imdb.com

1 Movie Recommendations

The Lost City of Z (2016) - Full Cast & Crew - IMDb










 **The Lost City of Z** (2016)
Full Cast & Crew

[IMDbPro](#) See agents for this cast & crew on IMDbPro

Directed by
James Gray ... (directed by)

Writing Credits
James Gray ... (written for the screen by)
David Grann ... (book by)

Cast (credits order) verified as complete

| | | | |
|---|------------------|-----|-----------------------|
|  | Charlie Hunnam | ... | Percy Fawcett |
|  | Robert Pattinson | ... | Henry Costin |
|  | Sienna Miller | ... | Nina Fawcett |
|  | Tom Holland | ... | Jack Fawcett |
|  | Edward Ashley | ... | Arthur Manley |
|  | Angus Macfadyen | ... | James Murray |
|  | Ian McDiarmid | ... | Sir George Goldie |
|  | Clive Francis | ... | Sir John Scott Keltie |
|  | Pedro Coello | ... | Tadju |

AMAZON ORIGINAL
CINDERELLA
WATCH WITH PRIME


The Lost City of Z


Details
[Full Cast and Crew](#)
[Release Dates](#)
[Official Sites](#)
[Company Credits](#)
[Filming & Production](#)
[Technical Specs](#)

[Explore More](#)

User Lists [Create a list »](#)

Related lists from IMDb users

 **films watched in 2021**
a list of 24 titles
created 8 months ago

 **Movies Based on Books**
a list of 39 titles
created 5 months ago

DATA COLLECTION

- Created link for movie cast and crew
- Scraped top five actors for each movie

LIMITATIONS

- Scraping took too long
- Only able to get ~50% of original dataset

METHODS & MODELING

- Features of Interest:
 - Movie description
 - Lead actor
 - Director rank
 - Top genre

Modeling

- Created CountVectorized matrix for movie description
- Added lead actor data to matrix/df
- Concatenated director data and genre data to matrix
- Created cosine similarity table based off this matrix
- Evaluated based off cosine similarity scores

RESULTS & CONCLUSIONS

- Results
 - Cosine Similarity scores improved from around 0.3 to 0.94 on test case
 - Movie recs seem to be fairly accurate
- Conclusions
 - Movie description, lead actor, genre, and director are good indicators of similarity

FUTURE RESEARCH

- Incorporate different features in the model
 - Language, country, year made, etc.
- Use pairwise distance, or another metric to evaluate the distance between movies
- Choose different types of models aside from ones that use CountVectorizer

**CHECK OUT THE
APP!**