CinFind

Your personalized movie recommender

Group: 07

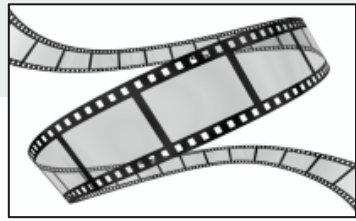
Team:

Arhab Hasan Khan
Balaji Udayakumar
Siddharth Khare

Index

Introduction

- Project aims to develop a movie recommendation tool that offers personalized suggestions to enhance user satisfaction
- Approach: Leveraging and applying cloud, big data and machine learning technologies for optimal suggestions
- We want to enhance user satisfaction by providing tailored movie recommendations
- How do movie reviews influence our decision to watch a movie? (Something to think about)



Research Questions

- What factors are more important in movie ratings: critical reviews (Tomatometer scores) or audience preferences?
- Can we create personalized movie recommendations based on individual user preferences?
- How accurate can predictive modeling be in figuring out movie ratings?
- Potential business outcomes of this project for streaming platforms and movie-aggregation websites

Dataset

1. **Source:** We sourced our dataset from Rotten Tomatoes, a renowned cinema review-aggregation website for films and television

2. **Fetched**: Utilizing web scraping techniques with BeautifulSoup and rottentomatoes python library, we extracted diverse movie data, including ratings, genres, and audience scores

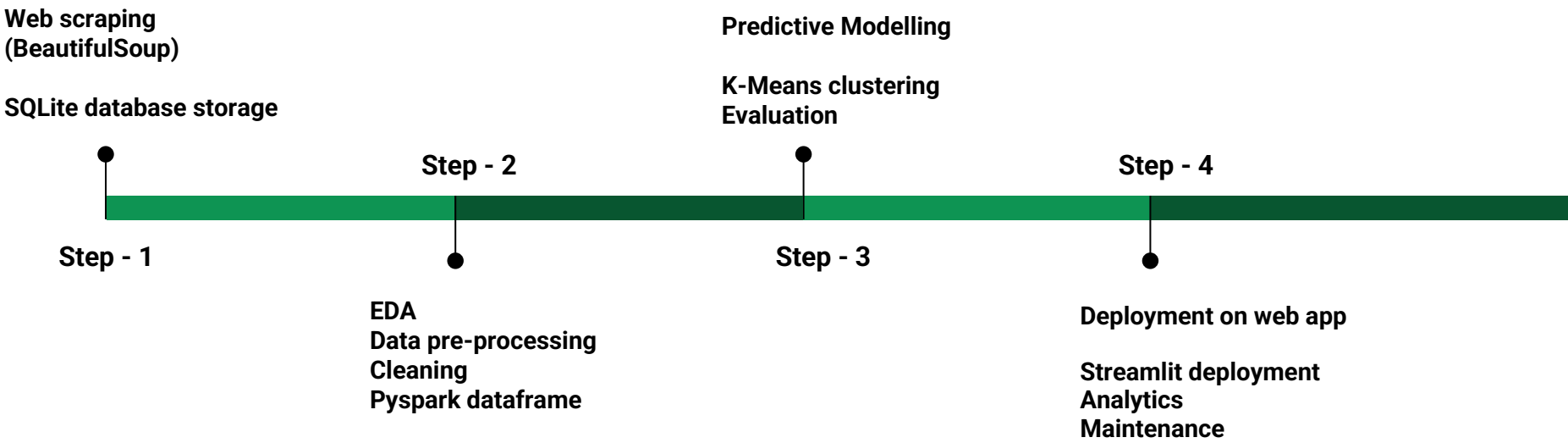
3. **Stored:** We stored the fetched data to a database using sqlite3 to get data quickly and easily.

4. **Processed**: The data is fetched from the DB and then essential columns such as Title, Tomatometer score, Audience score, Actors, Genre, Rating are processed to generate results from ML model.

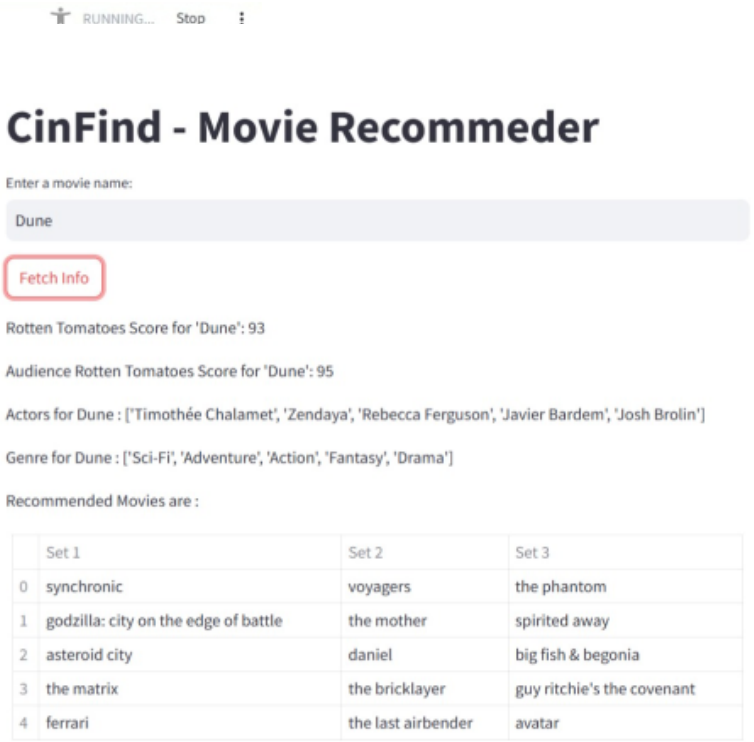
Methodology

Files needed to run the web app:

- movies.db
- app.py



Web Application



	Set 1	Set 2	Set 3
0	synchronic	voyagers	the phantom
1	godzilla: city on the edge of battle	the mother	spirited away
2	asteroid city	daniel	big fish & begonia
3	the matrix	the bricklayer	guy ritchie's the covenant
4	ferrari	the last airbender	avatar

Results and Findings

1. **Reviews matter**

- Both critical reviews (Tomatometer scores) and audience preferences significantly influence movie ratings
- Certain genres such as drama and sci-fi tend to have high-rated movies in our dataset

2. **Predictions:**

- The K-Means clustering model demonstrates promising cluster creation in estimating similar movies based on features like genre, rating, cast, and title
- Streaming platforms and movie-aggregation websites can leverage personalized recommendations to enhance user engagement and satisfaction.

Conclusions

These are the conclusions in our project:

1. Utilized web scraping, relational databases, machine learning, data pre-processing, and web application deployment for creating a personalized movie recommender

2. We shed some light on critical factors influencing movie ratings, genre trends, and the effectiveness of personalized recommendations.

3. **Future Opportunities:** Moving forward, there are a lot of opportunities for further research in this domain and we can further integrate this app to Pyspark Streaming and can compare and update(if not found) the records in the DB .

References

1. Rottentomatoes-python Library · PyPI
   https://pypi.org/project/rottentomatoes-python/

1. Localtunnel for Web Application Hosting
   https://github.com/localtunnel/localtunnel
   https://localtunnel.github.io/www/

1. “Machine Learning based Movie Recommendation System”
   https://hcis-journal.springeropen.com/articles/10.1186/s13673-018-0161-6

1. “An Efficient movie recommendation algorithm based on improved k-clique methods”
   https://www.academia.edu/51025889/A_Research_Paper_on_Machine_Learning_based_Movie_Recommendation_System

2. Other referenced links:
   https://www.reddit.com/r/webdev/comments/4649rw/rotten_tomatoes_api/
   https://www.rottentomatoes.com/browse/movies_at_home/genres

Thank You