COLLABORATIVE AND HYBRID RECOMMENDER SYSTEM

- By Team PyCoders

Submitted to: Professor Mohammad Saiful Islam

TEAM PYCODERS MEMBERS:

- Neelesh Vashist (C0858518)
- Rohit Kumar (C0859060)
- Mukul Bisht (C0857928)
- Saurabh Singh (C0859334)

PRODUCT INTRODUCTION

Welcome to our project, where we dive into the fascinating world of recommendation systems. Our main goal is to show you two powerful techniques: Collaborative Filtering and Hybrid Recommender Systems, made even more interesting by adding Sentiment Analysis.

Through this innovative approach, we aim to significantly enhance the user experience and engagement within the realm of movie recommendations. By combining user preferences and sentiments, our system provides personalized suggestions that resonate with each individual user.

Our team is excited to showcase the impact of this project and the ways it can revolutionize the way users discover and enjoy movies.

COLLABORATIVE FILTERING OVERVIEW

Collaborative Filtering is a technique that predicts user interests by collecting preferences from multiple users.

One key method within Collaborative Filtering is Singular Value Decomposition (SVD), which involves reducing dimensionality in user-item rating matrices.

The user-item rating matrix is pivotal, where each cell represents a user's rating for a movie.

Our approach utilizes matrix factorization and cosine similarity to make predictions.

This method benefits from leveraging collective user behavior.

COLLABORATIVE RECOMMENDER SYSTEM RESULTS

Our Collaborative Filtering model achieved an RMSE score of 0.5800.

RMSE (Root Mean Squared Error) is a critical metric for evaluating model accuracy. It quantifies the difference between predicted and actual ratings.

This score reflects the precision and effectiveness of our Collaborative Filtering technique in generating recommendations that align with user ratings.

TOP MOVIES PREDICTED BY COLLABORATIVE FILTERING

Collaborative filtering identifies top movie suggestions for users based on their past ratings.

These personalized recommendations reflect user preferences.

```
Top 10 Movies with Highest Predicted Ratings:
1. Star Wars (1977): 2.69
2. Fargo (1996): 2.21
3. Return of the Jedi (1983): 2.17
4. Contact (1997): 2.03
5. Raiders of the Lost Ark (1981): 1.89
6. Godfather, The (1972): 1.87
7. Toy Story (1995): 1.86
8. English Patient, The (1996): 1.85
9. Silence of the Lambs, The (1991): 1.78
10. Scream (1996): 1.71
```

```
10 Least Recommended Movies:
1. Very Natural Thing, A (1974): 0.00
2. War at Home, The (1996): 0.00
3. Sunchaser, The (1996): 0.00
4. King of New York (1990): 0.00
5. Office Killer (1997): 0.00
6. Walk in the Sun, A (1945): 0.00
7. Careful (1992): 0.00
8. Vermont Is For Lovers (1992): 0.00
9. Vie est belle, La (Life is Rosey) (1987): 0.00
10. Quartier Mozart (1992): 0.00
```

Figures for:

Top 10 Movies with Highest Predicted Ratings (left fig.), **10 Least Recommended Movies (right fig.)**

TOP MOVIES PREDICTED BY COLLABORATIVE FILTERING (CONT.)

Collaborative filtering identifies top movie suggestions for users based on their past ratings.

These personalized recommendations reflect user preferences.

```
Top 10 Recommended Movies for User (User 1 in this case):
1. English Patient, The (1996): 5.88
2. L.A. Confidential (1997): 4.20
3. Titanic (1997): 4.02
4. Sense and Sensibility (1995): 3.99
5. Secrets & Lies (1996): 3.55
6. In & Out (1997): 2.94
7. Evita (1996): 2.86
8. Emma (1996): 2.73
9. Apt Pupil (1998): 2.68
10. Air Force One (1997): 2.56
```

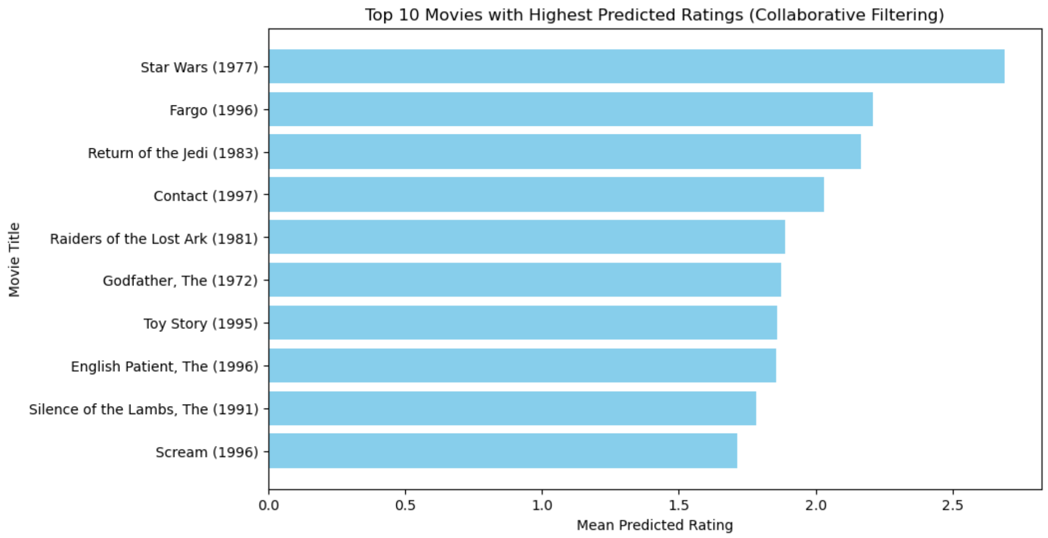
Figure for:

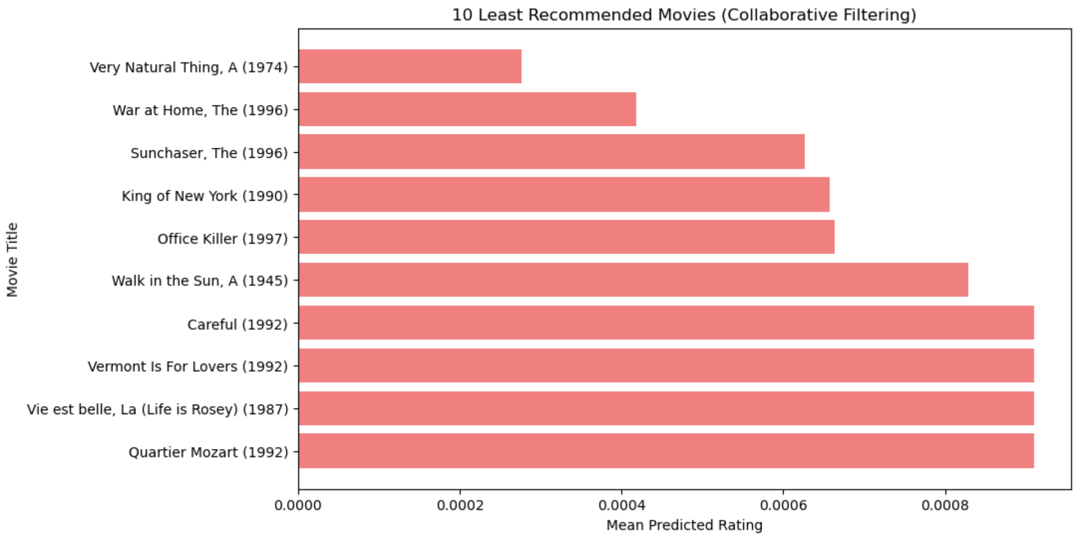
Top 10 Top Recommended Movies for User (User 1 in this case)

VISUALIZATION OF TOP RATED MOVIES

Figures for:

Top 10 Movies with Highest Predicted Ratings (left fig.), 10 Least Recommended Movies (right fig.)





HYBRID RECOMMENDER SYSTEM

Our Hybrid Recommender System seamlessly integrates Collaborative Filtering with Sentiment Analysis.

In addition to user ratings, we harness the power of Sentiment Analysis to incorporate insights from YouTube comments. By factoring in sentiment, we further refine recommendations to reflect not only numerical ratings but also the emotional resonance of movies with users.

DATA FETCHING: YOUTUBE VIDEO IDS AND COMMENTS

Data acquisition is a crucial aspect of our project. We employed the YouTube Data API to fetch essential information for our recommendation system.

Fetching YouTube Video IDs and Comments:

- We fetched YouTube video IDs and YouTube comments for movies using relevant search queries.
- Our system leveraged multiple API keys to ensure comprehensive data collection.

Our diligent data fetching process lays the foundation for a recommendation system that is both accurate and emotionally resonant.

SENTIMENT ANALYSIS OVERVIEW

Sentiment Analysis involves gauging emotions and opinions expressed in text. For this project, we are using TextBlob Library.

This method help us extract sentiments from YouTube comments related to each movie. By quantifying user sentiments, we gain valuable insights into the emotional connection between users and movies.

SENTIMENT ANALYSIS USING TEXTBLOB

- Sentiment analysis gauges viewer sentiments in YouTube comments.
- We employed TextBlob to compute sentiment polarity for each comment.
- Positive, negative, or neutral sentiments are quantified on a scale.
- This sentiment analysis adds a new dimension to our recommendation system.

	movie_id	movie_title	youtube_video_id	normalized_sentiment_score
0	1	Toy Story (1995)	v-PjgYDrg70	0.500000
1	2	GoldenEye (1995)	lcOqUE0u1LM	0.500000
2	3	Four Rooms (1995)	0hu12MP7b1U	1.000000
3	4	Get Shorty (1995)	t2QcA-KoF5s	0.875000
4	5	Copycat (1995)	5Pp64srMAH4	0.654167

Figure for:
Movies with Normalized Sentiment Scores using TextBlob
Note: We created a new dataframe for demonstration purpose.

HYBRID PREDICTIONS USING COLLABORATIVE AND SENTIMENT ANALYSIS

Our Hybrid Recommender System demonstrates an innovative combination of Collaborative Filtering and Sentiment Analysis, offering a holistic approach to movie recommendations.

Our Hybrid System thrives on integrating Sentiment Analysis into the recommendation process.

This fusion of user ratings and sentiments results in a comprehensive array of movie suggestions that not only resonate with user preferences but also tap into emotional connections, promising a richer and more personalized movie discovery journey.

TOP RECOMMENDED MOVIES USING HYBRID FILTERING

Leveraging hybrid scores, we've identified top 10 and bottom 10 recommended movies.

These movies cater to both user preferences and sentiments.

```
Top 10 Recommended Movies:
Movie: Scream of Stone (Schrei aus Stein) (1991), Hybrid Score: (1682, 'Scream of Stone (Schrei aus Stein) (1991)', 0.09)
Movie: You So Crazy (1994), Hybrid Score: (1681, 'You So Crazy (1994)', 0.19285714285714284)
Movie: Sliding Doors (1998), Hybrid Score: (1680, 'Sliding Doors (1998)', 0.15)
Movie: Sliding Doors (1998), Hybrid Score: (1680, 'Sliding Doors (1998)', 0.15)
Movie: B. Monkey (1998), Hybrid Score: (1679, 'B. Monkey (1998)', 0.15)
Movie: Mat' i syn (1997), Hybrid Score: (1678, "Mat' i syn (1997)", 0.14607142857142857)
Movie: Sweet Nothing (1995), Hybrid Score: (1677, 'Sweet Nothing (1995)', 0.15)
Movie: War at Home, The (1996), Hybrid Score: (1676, 'War at Home, The (1996)', 0.18)
Movie: Sunchaser, The (1996), Hybrid Score: (1675, 'Sunchaser, The (1996)', 0.15)
Movie: Mamma Roma (1962), Hybrid Score: (1674, 'Mamma Roma (1962)', 0.19166666666666665)
```

```
10 Least Recommended Movies:
Movie: Richard III (1995), Hybrid Score: (10, 'Richard III (1995)', 0.3)
Movie: Dead Man Walking (1995), Hybrid Score: (9, 'Dead Man Walking (1995)', 0.18)
Movie: Babe (1995), Hybrid Score: (8, 'Babe (1995)', 0.1425)
Movie: Twelve Monkeys (1995), Hybrid Score: (7, 'Twelve Monkeys (1995)', 0.15)
Movie: Shanghai Triad (Yao a yao yao dao waipo qiao) (1995), Hybrid Score: (6, 'Shanghai Triad (Yao a yao yao dao waipo qiao) (1995)', 0.15)
Movie: Copycat (1995), Hybrid Score: (5, 'Copycat (1995)', 0.19625)
Movie: Get Shorty (1995), Hybrid Score: (4, 'Get Shorty (1995)', 0.2625)
Movie: Four Rooms (1995), Hybrid Score: (3, 'Four Rooms (1995)', 0.3)
Movie: GoldenEye (1995), Hybrid Score: (2, 'GoldenEye (1995)', 0.15)
Movie: Toy Story (1995), Hybrid Score: (1, 'Toy Story (1995)', 0.15)
```

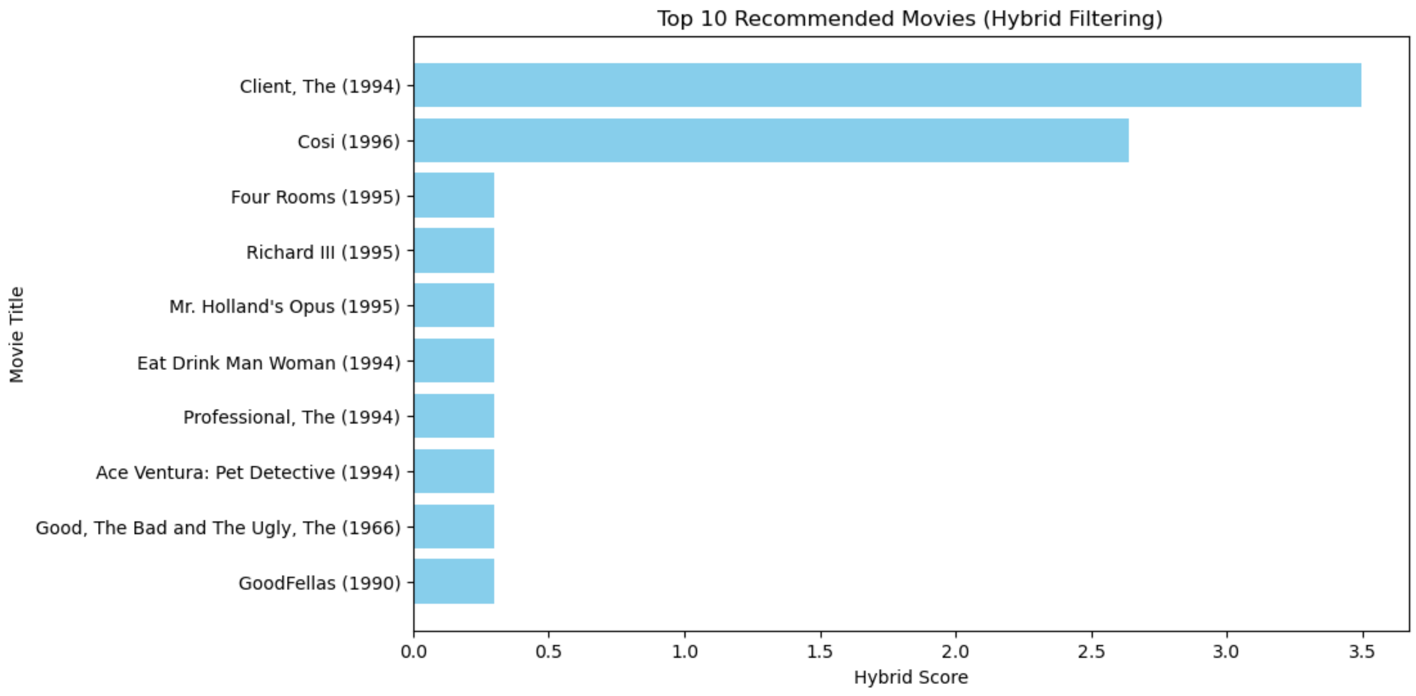
Figures for:

Top 10 Recommended Movies (top fig.), **10 Least Recommended Movies (bottom fig.)**

VISUALIZATION OF TOP RATED MOVIES

Figure for:

Top 10 Movies with Highest Predicted



CONCLUSION

In conclusion, our Collaborative and Hybrid Recommender System project has successfully demonstrated the potential of combining Collaborative Filtering with Sentiment Analysis.

Through this synergy, we've developed two types of recommendation systems that not only predicts user preferences based on ratings but also incorporates the emotional connection users share with movies, as revealed through sentiments.

Our project opens avenues for enhancing the way users discover and engage with movies, showcasing the significant role technology plays in personalizing and enriching user experiences.

THANK YOU

- Project by **Team PyCoders**

Team Members:

- Neelesh Vashist (C0858518)
- Rohit Kumar (C0859060)
- Mukul Bisht (C0857928)
- Saurabh Singh (C0859334)