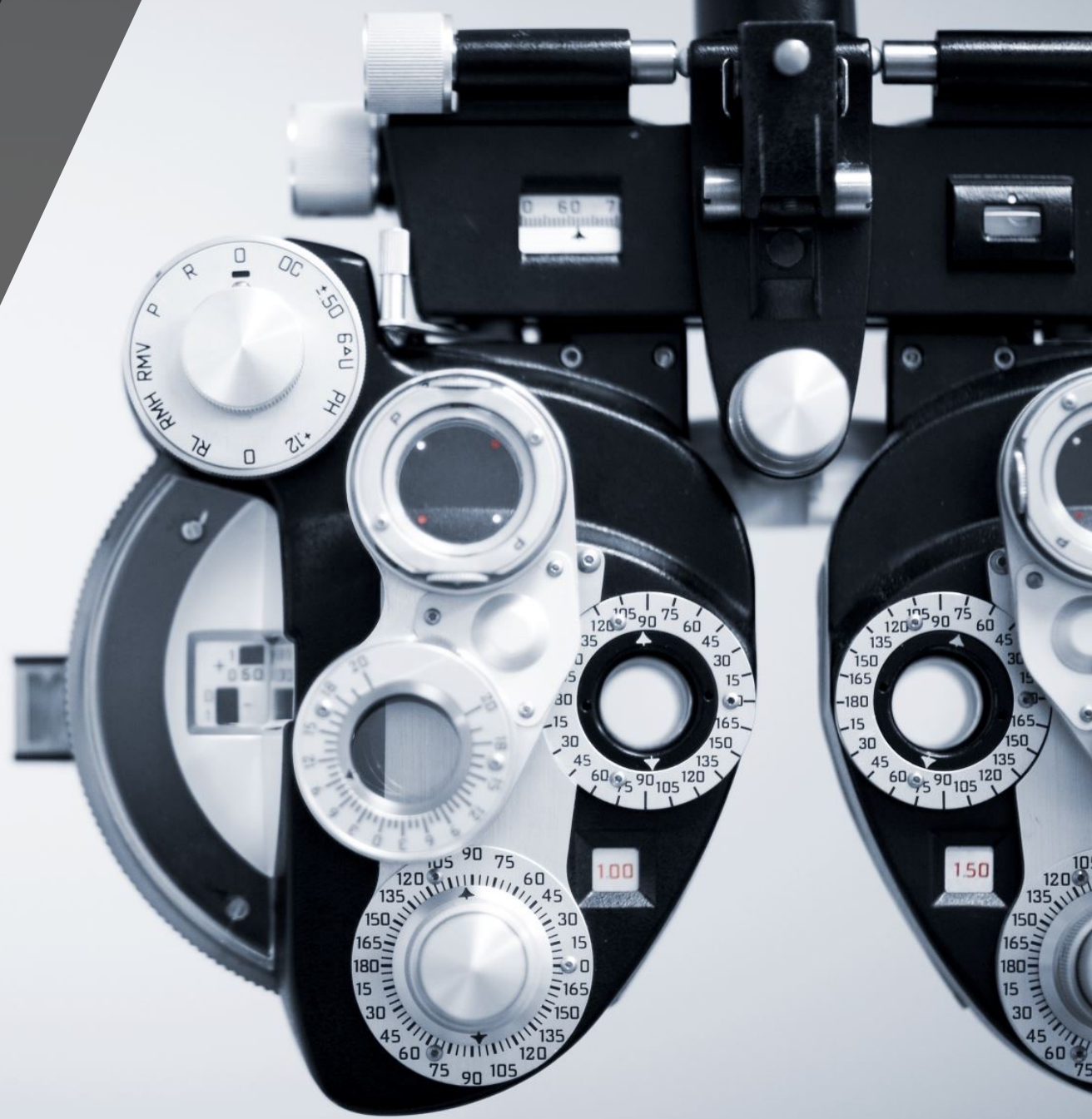


Movie Recommendation System

by Elena Burlando



Summary

Objective:

- to develop a Top-N movie recommendations for existing clients with an improved Cumulative Hit Rate (to ~ 0.05) for ratings ≥ 4.0 .

Projected Outcome:

- test multiple recommendation systems: popularity, content-based, Collaborative Filtering, and Matrix Factorization.
- recommend two-three systems with highest cHR.



outline



BUSINESS
PROBLEM



DATA AND
METHODS



RESULTS



CONCLUSIONS

BUSINESS PROBLEM



LOW CUSTOMER
ENGAGEMENT



CUMULATIVE HIT RATE IS
AT 0.01

DATA & METHODS

DATA

- Over 100K entries of movie ratings
- 9742 movies reviewed
- 610 users

FEATURES

- Genres, tags, date of release, links to movie details

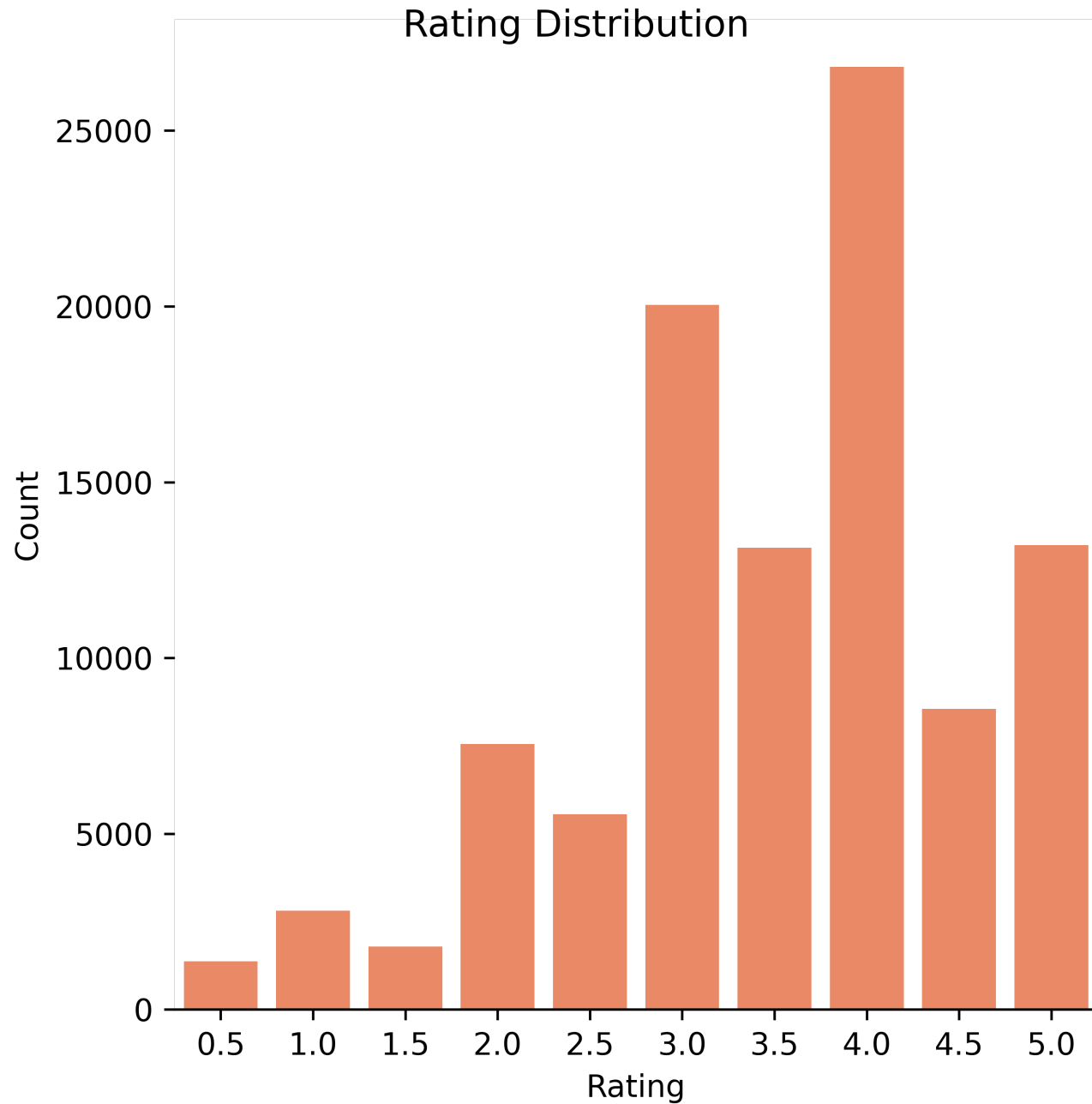
RECOMMENDATION SYSTEMS

- Popularity
- Content-based (genre)
- Collaborative Filtering
- Matrix Factorization (SDV, SVD++)

DATA SOURCES

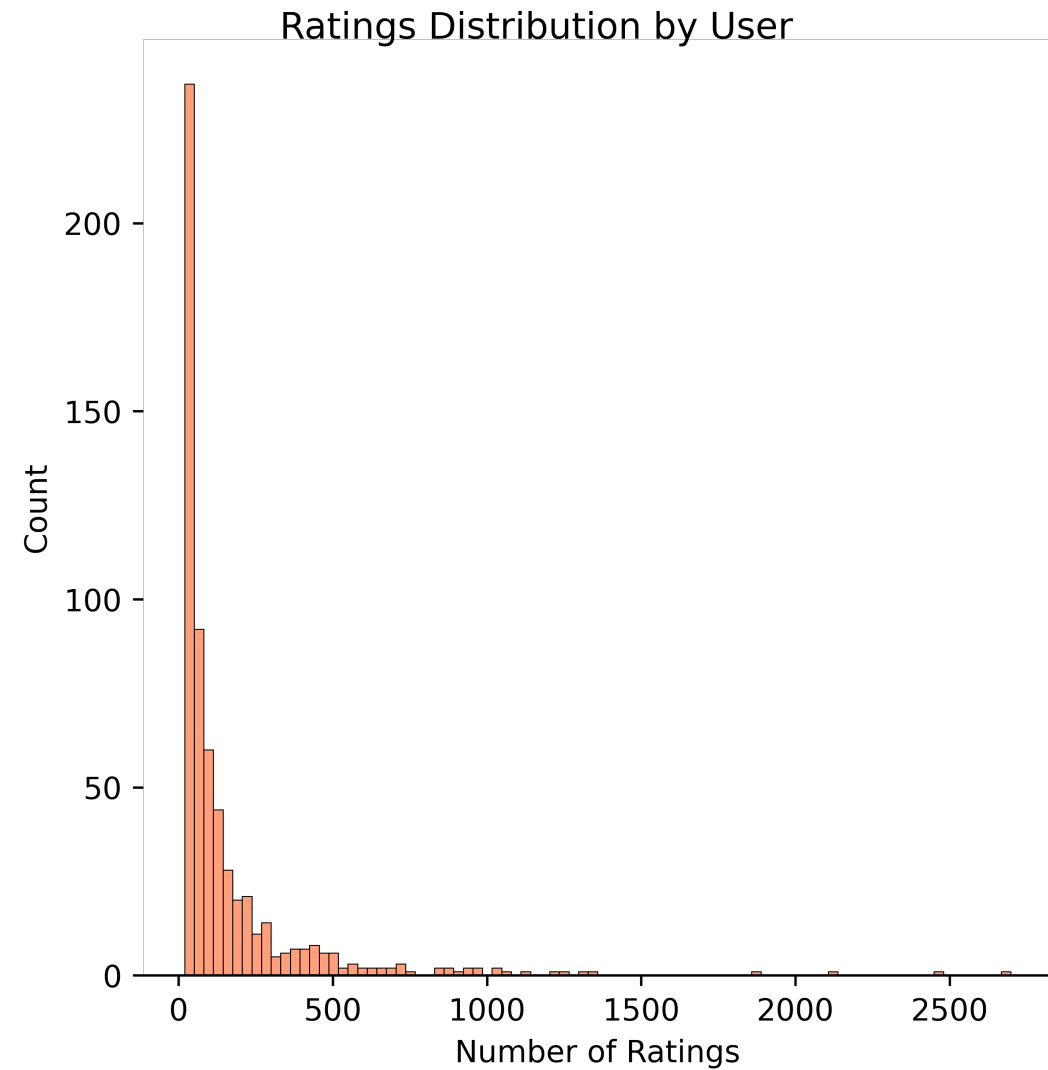
- [MovieLens small dataset](#)





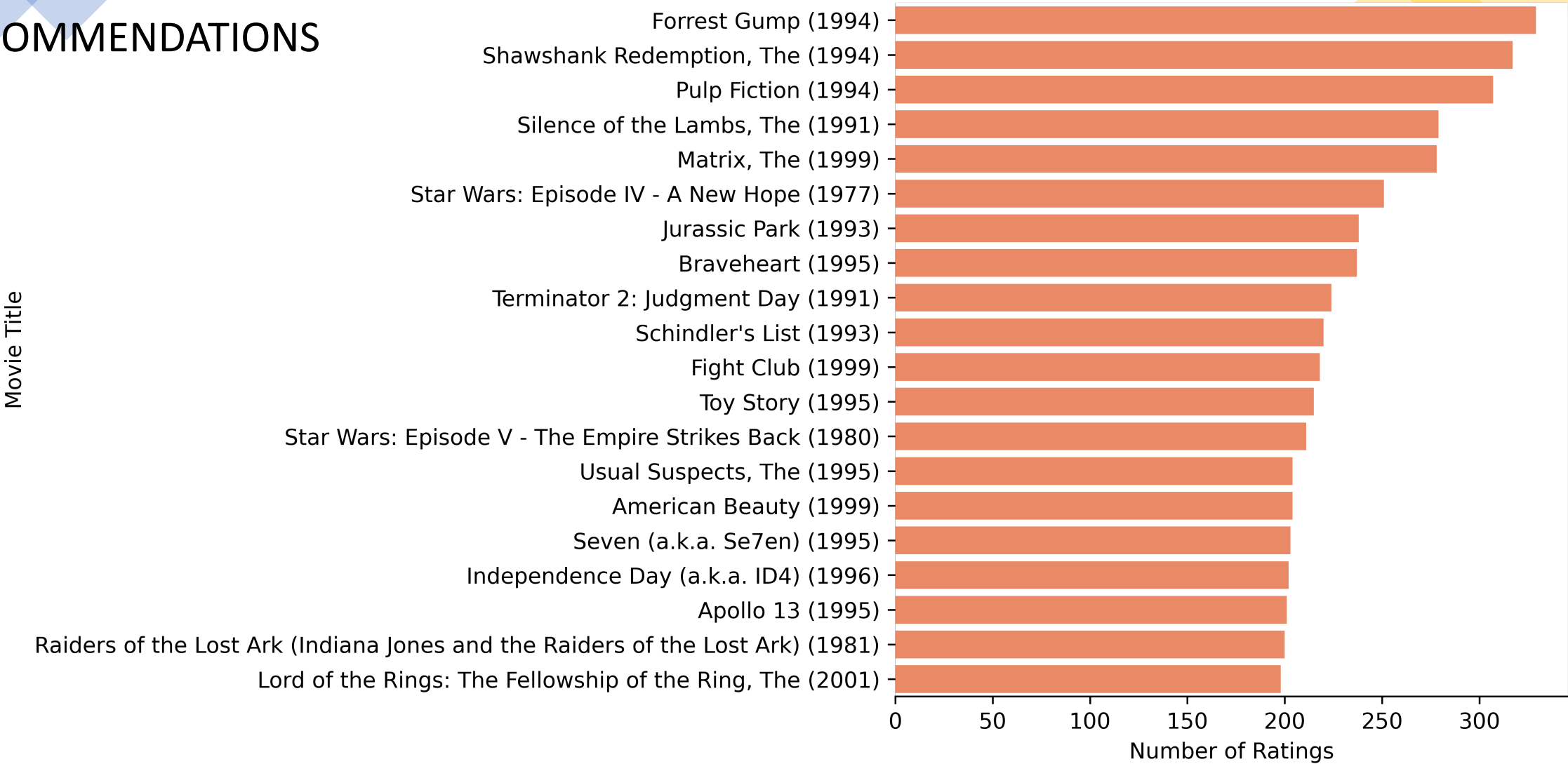
- Rating scale is 0-5.
- Rating distribution is skewed right indicating that customers tend to leave positive ratings.
- The largest # of rating given is 4.0 followed by 3.0, 3.5 and 5.

- Ratings distribution by user is skewed left indicating that majority of users have left 20+ reviews.
- Few users left 100+ reviews.



POPULARITY BASED RECOMMENDATIONS

Ratings Distribution by Movie



Untuned SVD

```
{'RMSE': 0.8806750857000041,  
'MAE': 0.6751593982768432,  
'HR': 0.03114754098360656,  
'cHR': 0.0425531914893617,  
'ARHR': 0.011357012750455373,  
'Coverage': 0.9245901639344263,  
'Diversity': 0.031900319933304444,  
'Novelty': 505.7342618384401}
```

We recommend:

Guess Who's Coming to Dinner (1967) 4.745014407216442

Three Billboards Outside Ebbing, Missouri (2017) 4.720015191777461

Paths of Glory (1957) 4.6446103691319545

Streetcar Named Desire, A (1951) 4.61991819185824

Godfather, The (1972) 4.612015465383254

Secrets & Lies (1996) 4.606239144936886

Adam's Rib (1949) 4.584542526932856

Trial, The (ProcÅ's, Le) (1962) 4.582209866985488

Neon Genesis Evangelion: The End of Evangelion (Shin seiki Evangelion GekijÅ-ban: Air/Magokoro wo, kimi ni) (1997)
4.577752288908924

Five Easy Pieces (1970) 4.56212527326592

Untuned SVD++

```
{'RMSE': 0.8681853914083167,  
'MAE': 0.6640591715627461,  
'HR': 0.03114754098360656,  
'cHR': 0.0425531914893617,  
'ARHR': 0.013036690085870413,  
'Coverage': 0.9147540983606557,  
'Diversity': 0.07063896593707153,  
'Novelty': 875.9580485952923}
```



We recommend:

Streetcar Named Desire, A (1951) 4.548908044586783

12 Angry Men (1957) 4.5399802872258395

Ran (1985) 4.515392734559682

Three Billboards Outside Ebbing, Missouri (2017) 4.514355757991221

Grand Day Out with Wallace and Gromit, A (1989) 4.50861699618624

Departed, The (2006) 4.501026716157305

High Noon (1952) 4.4935906701537105

Seventh Seal, The (Sjunde inseglet, Det) (1957) 4.475112667114752

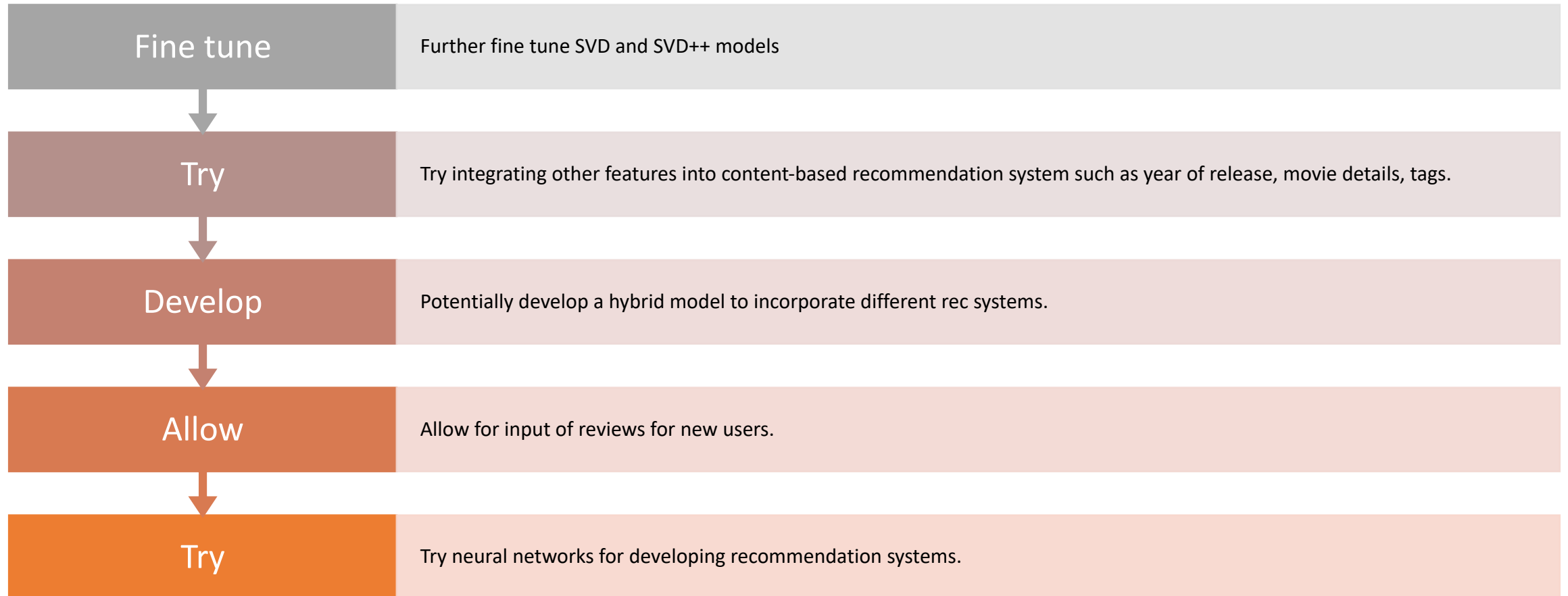
Secrets & Lies (1996) 4.46415371255078

Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb (1964) 4.4632009805588275

BUSINESS RECOMMENDATIONS

- Untuned SDV and SVD++ models have produced the best cHR at 0.042 and RMSE (SVD at 0.88 and SVD++ at 0.87) and MAE (SVD at 0.68 and SVD++ 0.66).
- SVD++ has output the best Top 10 recommendations with half of recommended movies that I would watch.

NEXT STEPS



Thank you!

[Link to Notebook](#)