



MOVIE RECOMMENDATIONS

Using Collaborative Filtering to implement an
item recommendation model

I don't know. What do you want to watch?

Austin B.

Overview



Goal



Data & Exploration



Modeling



Utility

Goal

Create a system to recommend movies to users

In: Movie Title **OR** Movie Ratings

Out: Users who liked this, also like:



DATA

MovieLens Data

- Data contains:
 - Users
 - Titles
 - Ratings
- Used 2 million ratings
- ~7K movies



DATA SCOPE

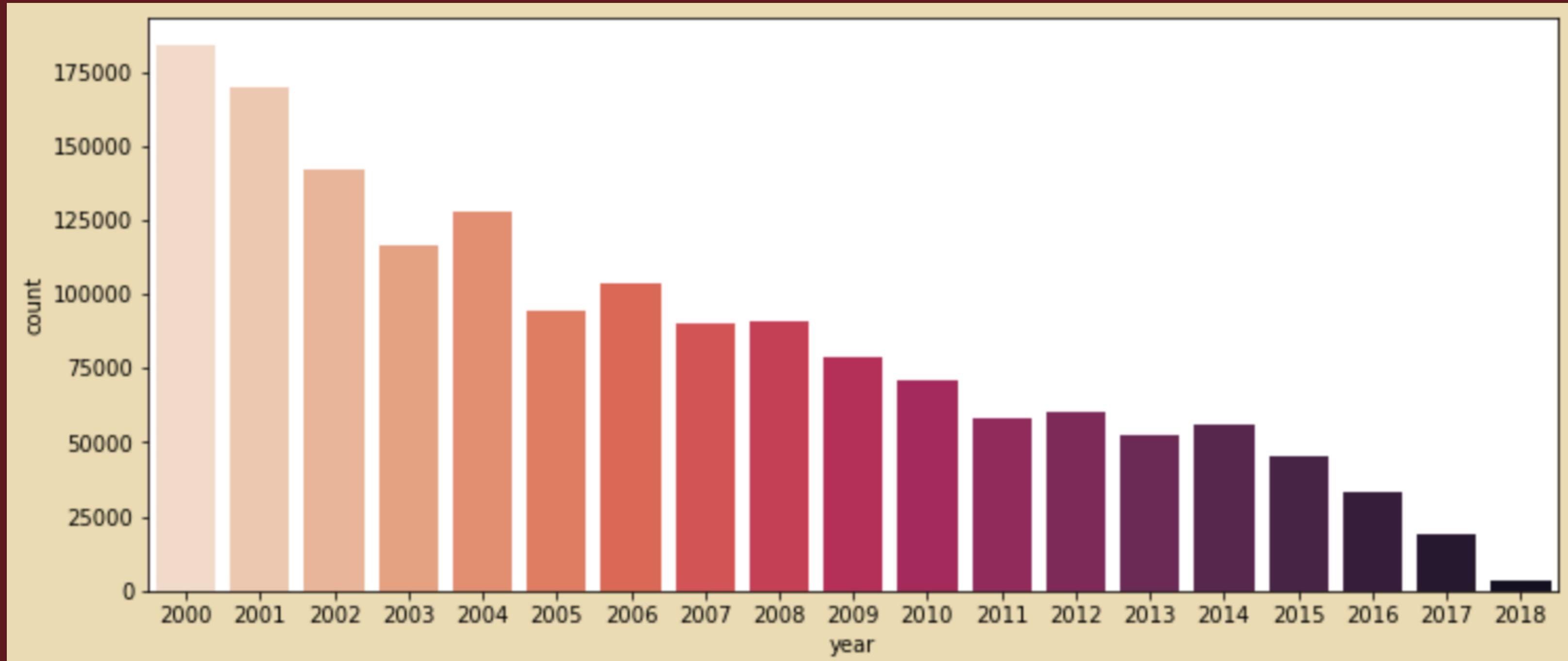
Recently released movies, excluding those with few to no ratings

- Cleaned missing and filler values
- Extracted release years from titles
- Used only movies between 2000-2018
- Filtered out movies with less than 25 ratings



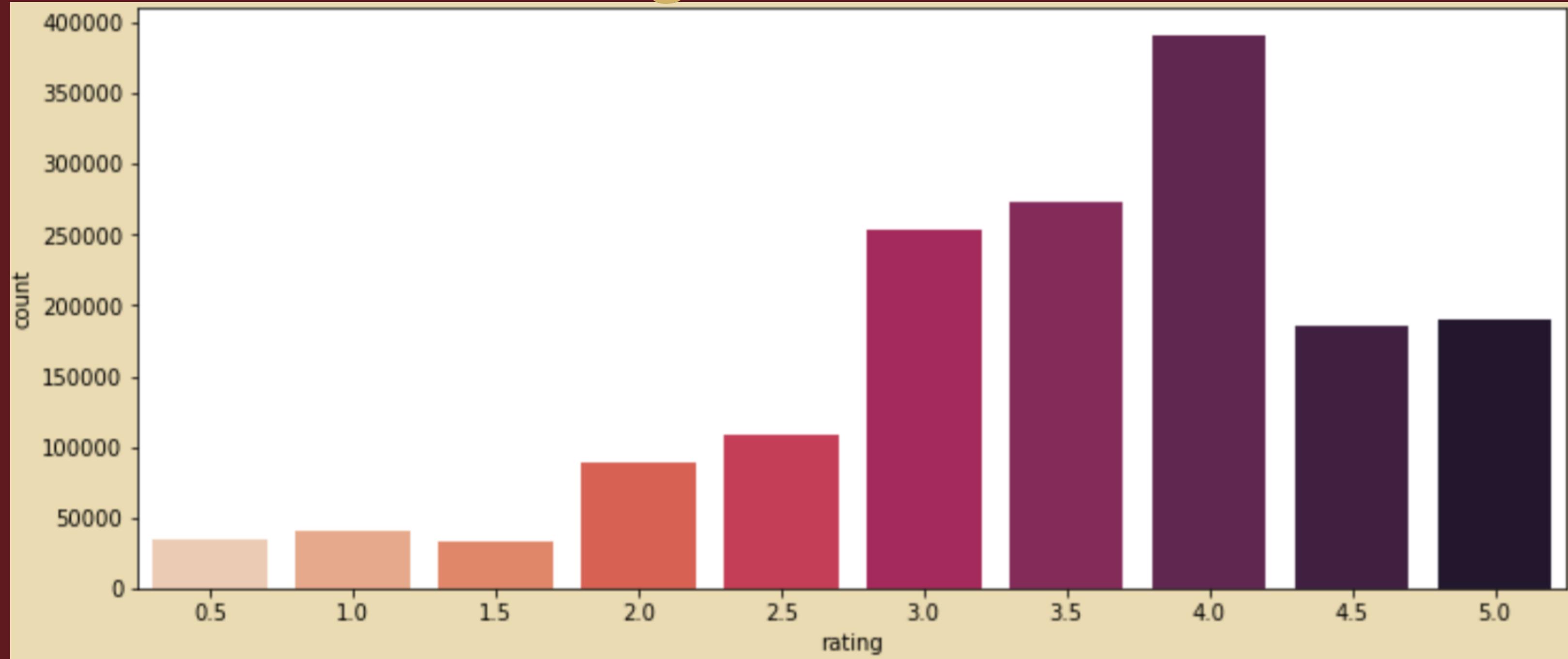
EXPLORATION

Ratings over years



EXPLORATION

Rating Distribution



EXPLORATION

20 Highest Rated Movies

But... have you SEEN Planet Earth

title	
Planet Earth II	4.510417
Planet Earth	4.478689
Bitter Lake	4.409091
Human	4.400000
Band of Brothers	4.376712
God Grew Tired of Us	4.375000
Sorry to Bother You	4.375000
Black Mirror	4.372470
Over the Garden Wall	4.372449
Like Stars on Earth	4.340278
Suicide Room	4.333333
AlphaGo	4.312500
Fresh Guacamole	4.312500
Between the Folds	4.307692
Sherlock	4.306122
Human Planet	4.296875
Century of the Self	4.287500
Puella Magi Madoka Magica the Movie Part II	4.277778
The Blue Planet	4.272727
Death on the Staircase	4.242424

MODELING

METRIC

Mean Absolute Error

An intuitive measure of the average error across all predicted ratings

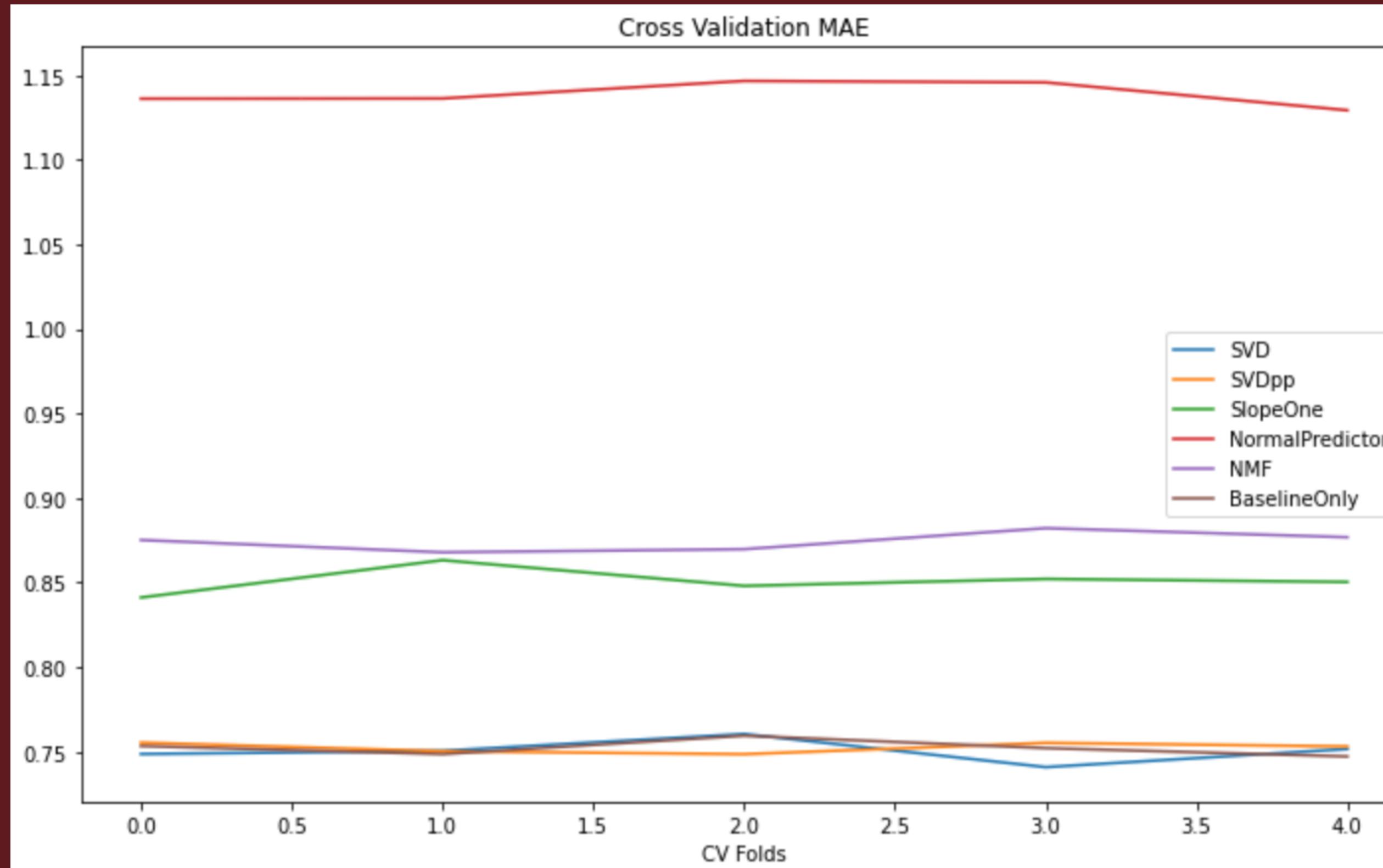
MODELS

- SVD
- SVDpp
- NMF
- SlopeOne
- NormalPredictor
- Baseline
- KnnWithZscore



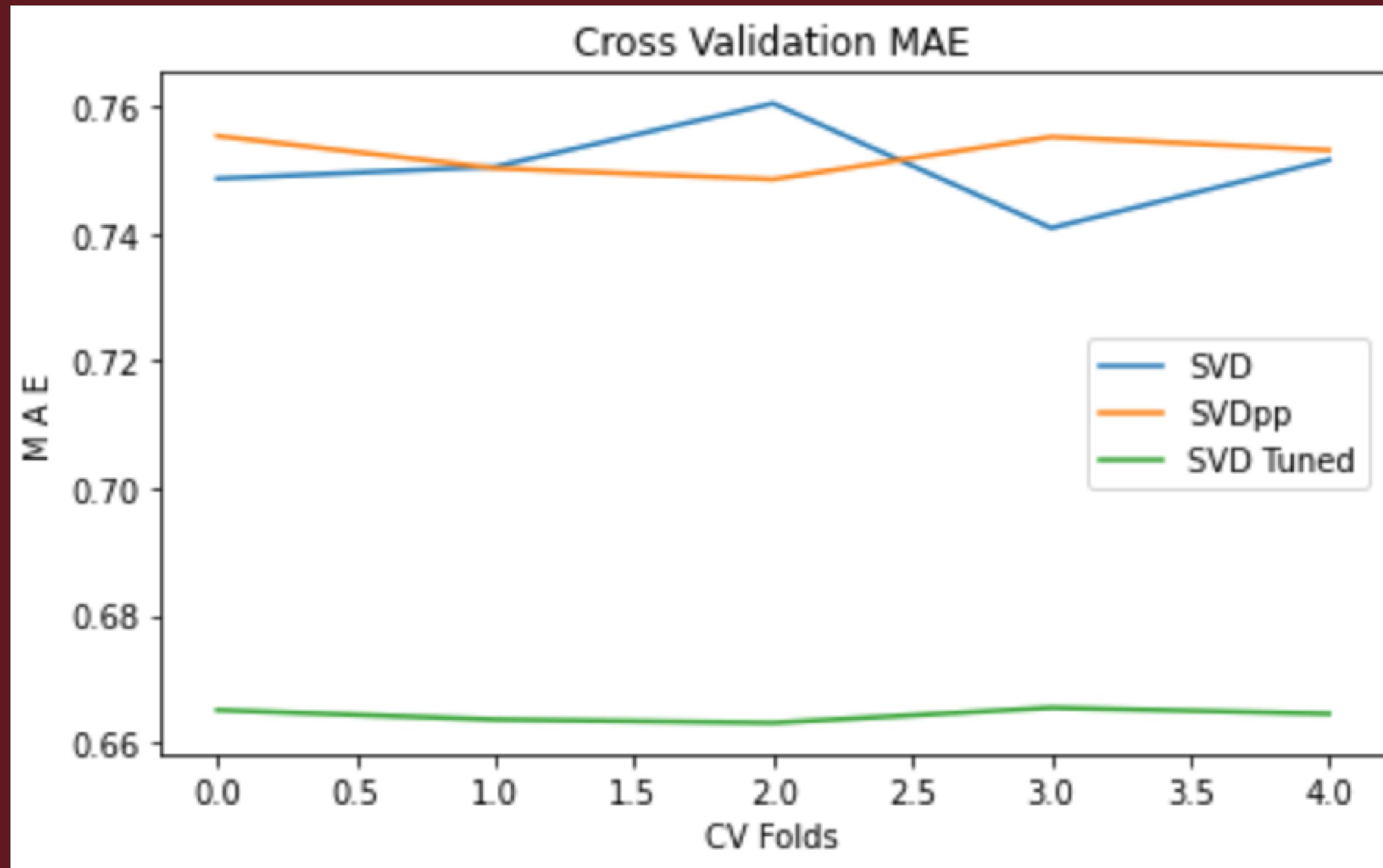
EXPLORATION & MODELING

Baseline Modeling



EXPLORATION & MODELING

Modeling



FINAL MODEL

MODEL

Support Vector Decomposition

Thank Netflix

Final Metrics

MAE: 0.6780

Pearson Correlation 0.57



RATINGS

.....	
295938	Django Unchained
Rating: 5	
1898523	Lord of the Rings
Rating: 4	
1472129	Larry the Cable Guy
Rating: 2	
1165090	Inception
Rating: 5	
621887	Pirate Radio
Rating: 4	
.....	



RECOMMENDATIONS

Movie Recs:

Recommendation # 1 : Planet Earth

Recommendation # 2 : Band of Brothers

Recommendation # 3 : Planet Earth II

Recommendation # 4 : Over the Garden Wall

Recommendation # 5 : Black Mirror

MOVIE IN, MOVIES OUT

Enter Movie Title... Spelled correctlyThe Lost City of Z

['Lord of the Rings', 'Monster', 'Silent Hill ', 'Cloverfield ', 'Suicide Squad ']

(169900, 'The Lost City of Z ')

Enter Movie Title... Spelled correctlyNight at the Museum

['Master and Commander', 'Brokeback Mountain ', 'Cloverfield ', 'The Accountant ', 'Guardian']

(119155, 'Night at the Museum')

IMPROVEMENTS

- Tuning model weights to smooth distribution bias
- Creating weighted hybrid model to incorporate content based filtering



THANK YOU



<https://github.com/Auustiino>