Phase 4 Project

Heath Rittler

Business Problem

Description: You are working for an **online streaming platform that offers** a vast collection of **movies and TV shows** to its subscribers. The platform wants to improve their movie recommendation system to enhance user engagement, increase user satisfaction, and ultimately drive more subscriptions.

Goal: **Build a recommendation system** that provides personalized movie recommendations to each user based on their preferences and viewing history.

Data

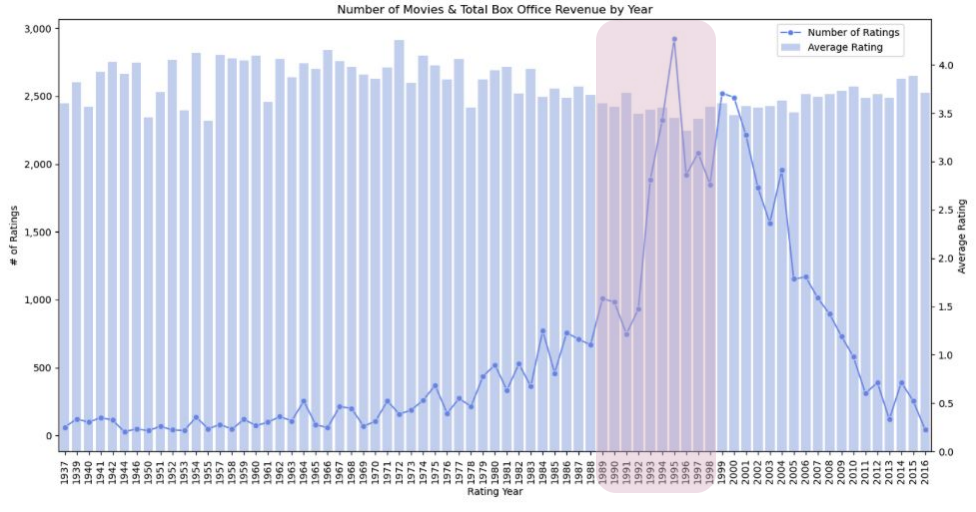
41k ratings *from*

450 movies *and*

600 users *with a rating*

scale *of* **0-5**

Ratings Over Time



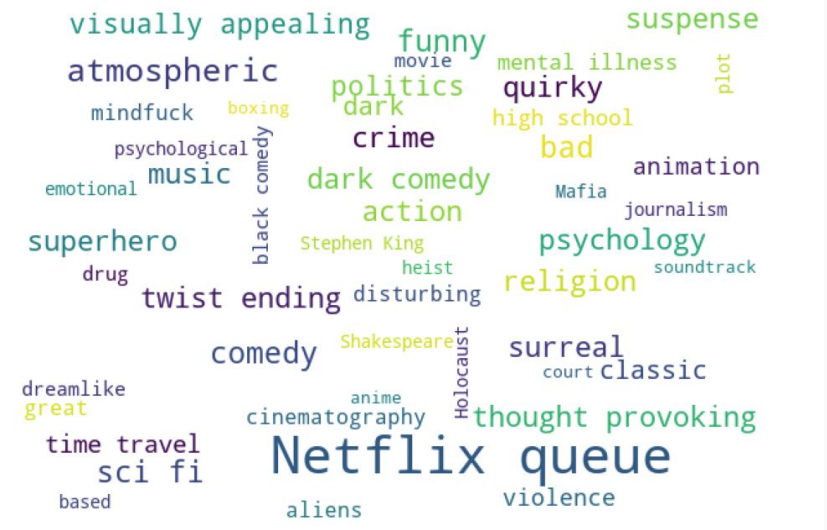
Approach & Goals

Collaborative based system

Content based system

User Input to seed initial interest

Movie Tags



Final Models

Primary metric is **RMSE**

Tuned **SVD** Model

Predicted ratings are off by approximately **0.80** units from actual

Generally **good** performance

Model Performance

model	RMSE	MAE	CV
svd_tuned	0.811943	0.619905	0.806629
baseline_svd	0.832547	0.639260	0.827359
knnmeans	0.837140	0.642421	0.830563
knnbasic	0.911832	0.702170	0.911136
nmf	0.948063	0.711590	0.940486

Content Recommender

Considering **genre** as our content filter - cosine similarity

Decent/ not great; no real **awareness**

Opportunity to include **tags or** movie **descriptions** to enhance

Content Recommendations for Toy Story

```
Similar movies to movie 1 :
1 Toy Story (1995)
3114 Toy Story 2 (1999)
4016 Emperor's New Groove, The (2000)
4886 Monsters, Inc. (2001)
2294 Antz (1998)
```

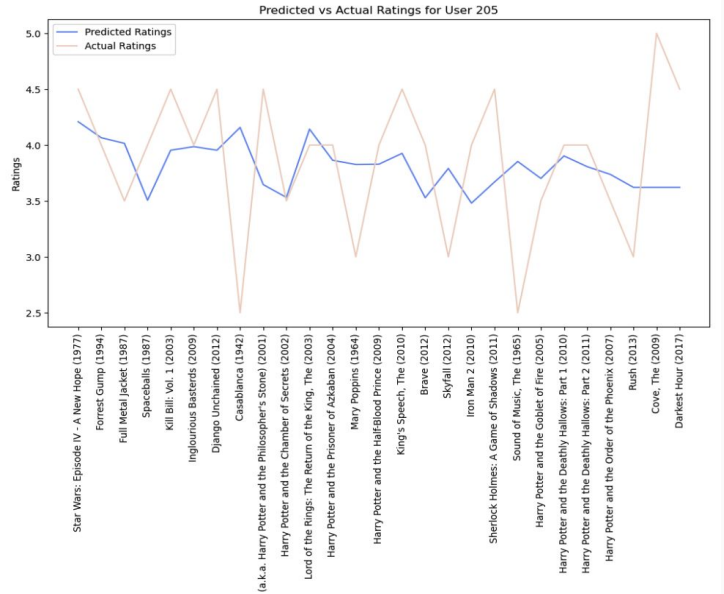
Collaborative Recommender

Rating **comparison** for sample user

Good predictions, when considering **RMSE of .80** (slide 5)

Look at **larger datasets**, and potentially online compute resources

Example User



Recommendations

A/B testing between users without the recommendations, and users with the recommendations to measure engagement.

Continue to evaluate content recommendation model to **improve hybrid approach** (tags/ descriptions)

Work with software dev/engineering to **implement** initial **rating capture** for new users

Consider **larger datasets** and **better compute** resources

Thank you!

Email: hrittler@gmail.com

Github: @heathlikethecandybar

LinkedIn: linkedin.com/in/heathrittler