

viiiiiiiiii / Flatiron_Phase4_Final_Vi_Bui Public

Phase 4 Project - Recommendation System

☆ 0 stars ⚡ 0 forks

Star

Unwatch

Code

Issues

Pull requests

Actions

Projects

Wiki

Security

main

...



viiiiiiiiii Add files via upload ...

now

23

[View code](#)

README.md



Flatiron Phase 4 - Final Project by Vi Bui

Movie Recommendation System



Overview

Vi(sion) Studios, a new streaming service, is looking to launch a concept called "Digital Cinema Night" where customers can build a "Cinema Night" around specific movies. They've hired us to build a Recommendation System in order to launch this concept.

Data, Methodology, and Analysis

Data source: Popular MovieLens dataset which captures Movies, Ratings, Genres, and Year.

More about the data from MovieLens: "Users were selected at random for inclusion. All selected users had rated at least 20 movies. No demographic information is included. Each user is represented by an id, and no other information is provided."

There are 9724 unique entries and 610 unique users in the data.

Models Built:

1. Singular Vector Decomposition Model (SVD)
2. k-Nearest Neighbor Baseline Model (KNNB)

3. Non-Negative Matrix Factorization Model (NMF)

Features included and created:

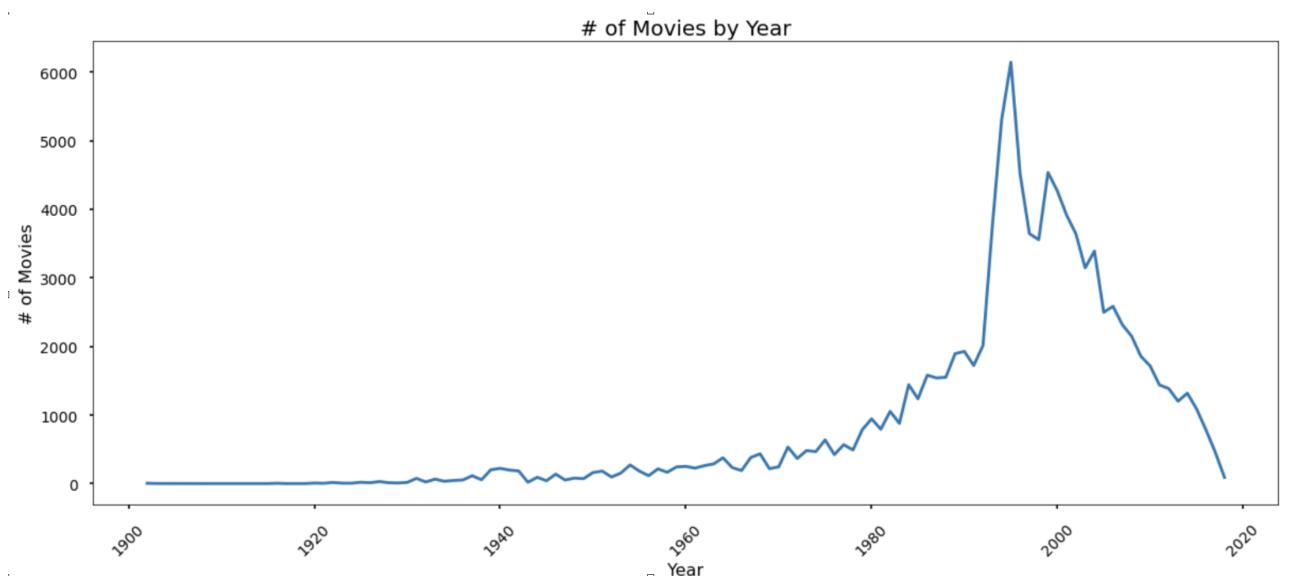
1. Movie ID (included)
2. User ID (included)
3. Ratings (included)
4. Year (cleaned & created)
5. Genre (cleaned & created)
6. Ratings Count (created)

BUSINESS VALUE

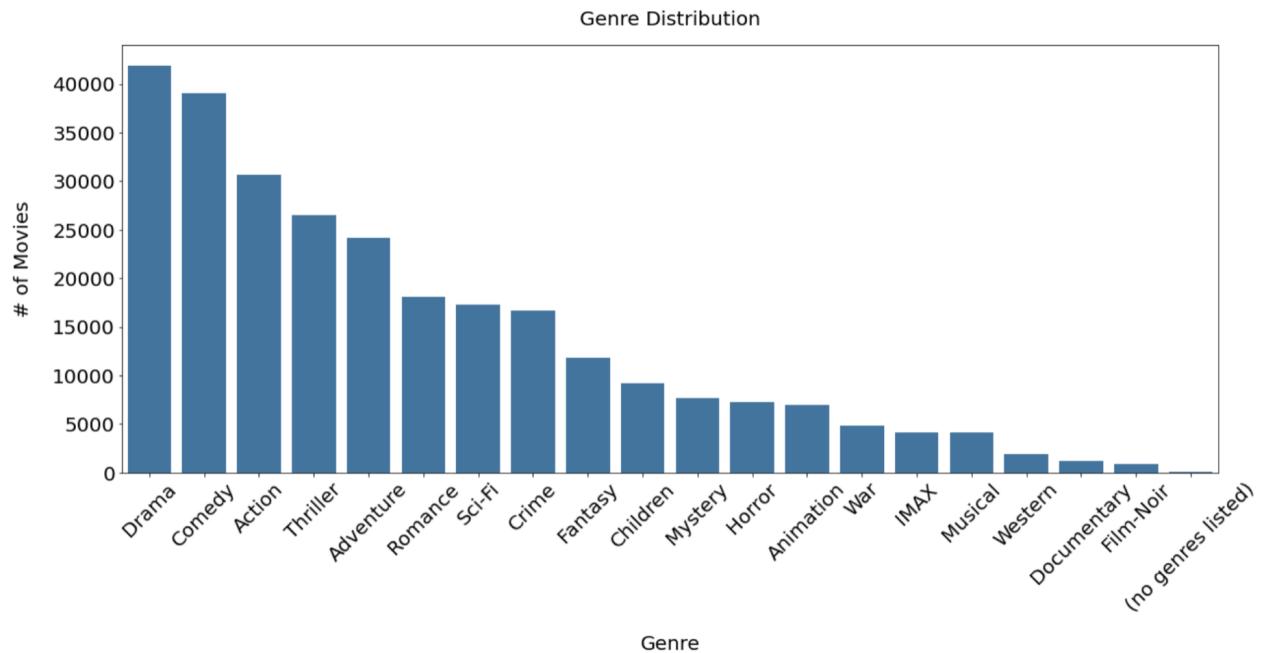
“Digital Cinema Night” will bring unique value to their users, and build long-term loyalty, by offering an experience for their users around a movie of choice vs. what users currently do (add items to their queue or have services automatically play movies without knowledge of recommendations)

OBSERVATIONS ABOUT THE DATA

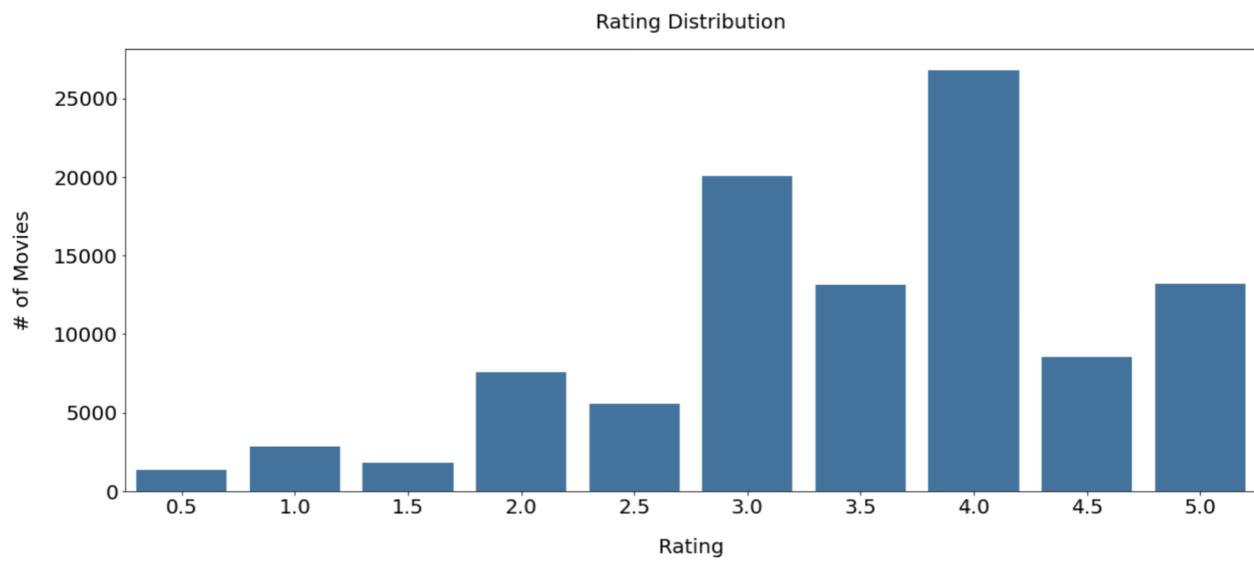
Most movies are from the 1990s to the early 2000s (key Millennial time period)



The most popular genres: Drama, Comedy, and Action



Users tend to rate in whole numbers and 4 (out of 5) is the most popular rating



Models & Metrics

We ran three models and used RMSE (root mean square error), which tells us how close our model's predicted

ratings are to actual ratings.

While all models showed strong results and predictions, we chose the Singular Value Decomposition (SVD) model, which had an RMSE of 0.86, meaning our predictions, on average, were 0.86 points away from actual ratings.

Singular Vector Decomposition Model (SVD)	k-Nearest Neighbor Baseline Model (KNNB)	Non-Negative Matrix Factorization Model (NMF)
* Pre-Tuned RMSE: 0.95 * Tuned RMSE: 0.86	* Pre-Tuned RMSE: 0.96 * Tuned RMSE: 0.39	* Pre-Tuned RMSE: 1.12 * Tuned RMSE: 0.30

RESULTS & RECOMMENDATIONS

Summary of recommendations

Use SVD Model to launch

- While all models showed improvement after tuning and KNN Baseline and NMF Models have low RMSEs (predictions are close to actual ratings), in order to reduce the risk of over-fitting, the SVD Model - with a strong RMSE of 0.86 - is recommended

Test for 6 months and re-evaluate

- Revisit KNN Baseline and/or NMF Models if necessary

Launch concept with Millennial (or Millennial-enthusiast) audience

- Movies leaned toward 1990s to early 2000s
- Opportunity to build early loyalty with an influential demographic

FUTURE WORK

- Build functionality around genre and year choices
- Extend audience beyond initial millennial focus
- Expand dataset with more movie choices & continually refine "Digital Cinema Night" system to make stronger recommendations to users
- Build exquisite user interface

BRINGING THE VI(SION) TO LIFE

DIGITAL CINEMA NIGHT!!

1. Users Input Movie

```
cinema_night_for('Toy Story')
```

2. Digital Cinema Night is created!

<code>cinema_night_for('Toy Story')</code>	<code>cinema_night_for('Shawshank Redemption, The')</code>	<code>cinema_night_for('Godfather, The')</code>
Incredibles, The 0.643301 125	Four Weddings and a Funeral Schindler's List	Godfather: Part II, The Schindler's List
Finding Nemo 0.618701 141	Usual Suspects, The	Fight Club
Aladdin 0.611892 183	Ocean's Eleven	Saving Private Ryan
Monsters, Inc. 0.490231 132	Green Mile, The	Goodfellas
Mrs. Doubtfire 0.446261 144		

Repository Structure

- Images
- README.me
- Vi_Bui_Phase4_Project_FINAL_Surprise.ipynb
- Vi_Bui_Phase4_Project_FINAL_Jupyter Notebook.pdf
- Vi_Bui_Phase4_Project_FINAL_Presentation.pdf
- Flatiron_Phase4_Final_Vi_Bui/edit/main/README.md.pdf

Releases

No releases published

[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Languages

- Jupyter Notebook 100.0%