

Business Problem

Netflix is looking to improve their recomendation system for new users. As part of a new trial membership program Netflix is looking to maximize their customer retention by providing the best possible recomendations. It is important that this new recomendation system will also improve recomendations for existing customers.

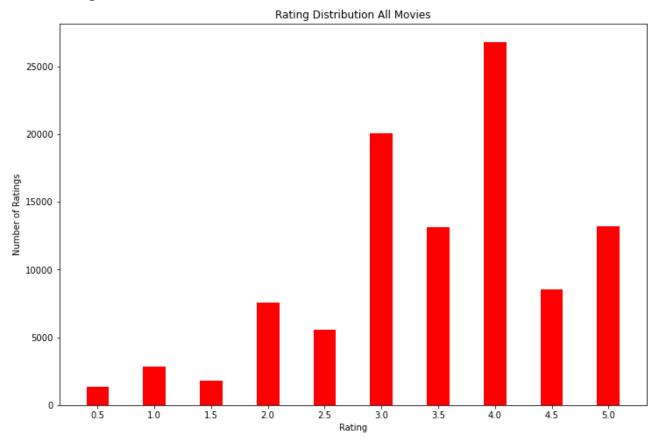
Data Used

data from MovieLens can be found in the 'data' folder

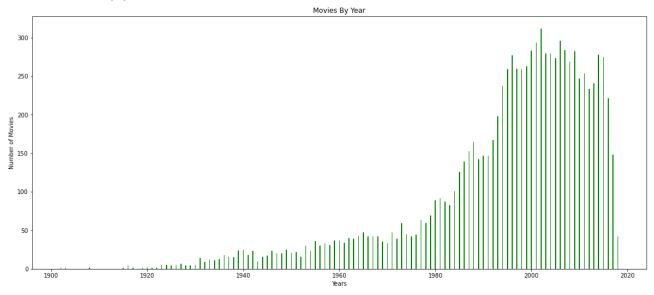
- movies.csv
- · ratings.csv
- tags.csv
- links.csv

movies.csv and ratings.csv were used to create the model and make predictions

Movie rating distribution can be seen below.



Movies used by year



Final Model

The final model was a Singular Value Decomposition(SVD) with hyper-tuned parameters using surprise library.

Model Evaluation

the metric used to evaluate the models was root-mean-square error (RMSE). Our best model had an RMSE of about .860.

Sample Predictions

Movie predictions from model. The top 5 rated movies.

STANDARD RECOMENDATIONS

Title: Billabong Odyssey (2003)

Genre: Documentary Projected Rating: 4.5

Title: How the Grinch Stole Christmas! (1966)

Genre: Animation | Comedy | Fantasy | Musical

Projected Rating: 4.5

Title: My Best Friend's Wedding (1997)

Genre: Comedy | Romance Projected Rating: 4.5

Title: I Love Trouble (1994)

Genre: Action | Comedy Projected Rating: 4.5

Title: Marat/Sade (1966)

Genre: Drama | Musical Projected Rating: 4.4

Filtered movie predictions. The top 5 rated movies after results have been fileterd by user survey.

YOUR CUSTOMIZED RECCOMENDATIONS

Title: Education, An (2009)

Genre: Drama | Romance Projected Rating: 4.2

Title: Despicable Me (2010)

Genre: Animation | Children | Comedy | Crime

Projected Rating: 4.2

Title: White Ribbon, The (Das weiße Band) (2009)

Genre: Drama | Mystery Projected Rating: 4.2

Title: Pirate Radio (2009)

Genre: Comedy | Drama Projected Rating: 4.2

Title: Gamer (2009)

Genre: Action|Sci-Fi|Thriller

Projected Rating: 4.2

Limitations

Our model will only be as good as our data. We should anticipate and expect improved metrics as our data increases.

Recommendations and Next Steps

- Develop plan to encourage users to rate more movies.
- Create new models using user login demographics.
- Use Natural Language Processing(NLP) on movie scripts to gain more insight on our users preferences.

Conclusion

After a thorough iterative modeling process, we determined that the hypertuned svd model would be the best selection at this point. Reccomendations would be made for the new trial users based on their initial reccomendation and the optional survey that limits the years of release and the genres of the recommendations.

Deliverables

- Non-Technical Presentation
- GitHub Repository
- Jupyter Notebook

Releases

No releases published Create a new release

Packages

No packages published Publish your first package

Languages

Jupyter Notebook 100.0%