

Movie Recommendations

Developing and Deploying a Recommendation Algorithm

Carlos Garza

Business Problem

- A new streaming company called ML Movies wants to implement an active movie recommendation system for its users
- Using a list of available films that have previously been rated by other users, develop a recommendation algorithm that generates curated movie recommendations.

Data Utilized

- The data for this project is sourced from [MovieLens](#). The data describes
 - 100,936 ratings
 - 9742 movies
 - 610 users

Baseline Model

- Surprise Library used to develop baseline models for three algorithms
- Hyper-parameter settings left as default for baseline testing
- Models cross validated, Average RMSE shown in table to the right.

Model	RMSE
SVD	0.873
KNNWithMeans	0.898
KNNBasic	0.947

Tuning and Results

- Using a brute force grid search that tries many combinations of model hyper-parameters, the SVD recommendation model's performance was slightly improved.
- The optimized model's performance is summarized in the table to the right.
- A script for a recommendation algorithm was created that requires no specialized knowledge to use.

Metric	Average Value
RMSE	0.856
MAE	0.656
Fit Time (s)	6.54

Deployment

- Model hyper-parameters are used in a python script algorithm

```
(learn-env) carlosgarza@x86_64-apple-darwin13 deployment % python3 recommendation.py

Welcome to ML Movies!

Let us find your new favorite movie!

Limit your search to a specific genre? (y/[n]) y

AVAILABLE GENRES:
['Adventure', 'Animation', 'Children', 'Comedy', 'Fantasy', 'Romance', 'Drama', 'Action', 'Crime', 'Thriller',
'Horror', 'Mystery', 'Sci-Fi', 'War', 'Musical', 'Documentary', 'IMAX', 'Western', 'Film-Noir']

From the list above, what genre do you prefer? Action
-----

To calculate the best movies for you,
we will need to hear what you think of some movies
you have already seen.

-----

How many movies would you like to rate for the algorithm? 3

6257    Miami Vice (2006)
Name: title, dtype: object
How do you rate this movie on a scale of 1-10, press n if you have not seen :
8

2029    Wild Wild West (1999)
Name: title, dtype: object
How do you rate this movie on a scale of 1-10, press n if you have not seen :
n

2898    Art of War, The (2000)
Name: title, dtype: object
How do you rate this movie on a scale of 1-10, press n if you have not seen :
6

6037    Legend of Zorro, The (2005)
Name: title, dtype: object
How do you rate this movie on a scale of 1-10, press n if you have not seen :
9

working....

Success!

How many movie recommendations would you like to see? 3

recommendation # 1
predicted movie rating: 9.66
Title: 5621    Neon Genesis Evangelion: The End of Evangelion...
Name: title, dtype: object

recommendation # 2
predicted movie rating: 9.35
Title: 2226    Fight Club (1999)
Name: title, dtype: object

recommendation # 3
predicted movie rating: 9.21
Title: 2283    Yojimbo (1961)
Name: title, dtype: object
(learn-env) carlosgarza@x86_64-apple-darwin13 deployment %
```

Conclusions

- Using singular value decomposition, the model created has a RMSE of 0.85
- The python script file using the SVD model provides the flexibility to filter by genre and is intuitive enough for a non technical audience
- The algorithm consistently finds movies to recommend the user with predicted user ratings of $>8/10$

Future Work

- Create GUI for a user to interact with the algorithm
- Code for the possibility to select more than one genre
- Create more robust code that is more flexible with user input
- Create a way to save recommendations or save and update a user profile

Thank You

Carlos Garza