

Your Anime Match Maker

Sherry Duong



[linkedin.com/in/sherryduong](https://www.linkedin.com/in/sherryduong)



github.com/sherryduong93



sherryduong@gmail.com



The Problem:

Finding a new anime takes a lot of time
searching through forums/websites

The Goal:

Streamline the process for finding the
next best anime to watch.

How?

Use existing data of anime and user ratings to build a Content Based & Collaborative Filter Recommender.

The Data:

Anime Dataset: 12K Animes

Anime Meta Data: 14K animes

Rating Dataset: 7M reviews on 11K
animes from 72K users

**Datasets are from Kaggle, scraped from MyAnimeList.net*



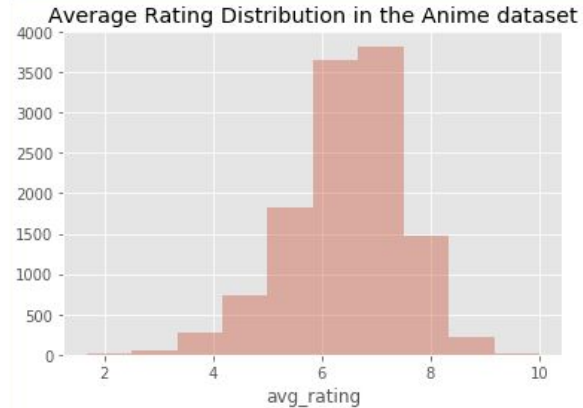
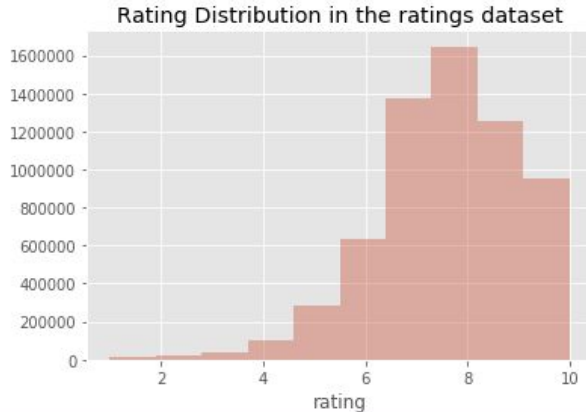
The Scope

How will I evaluate my efforts?

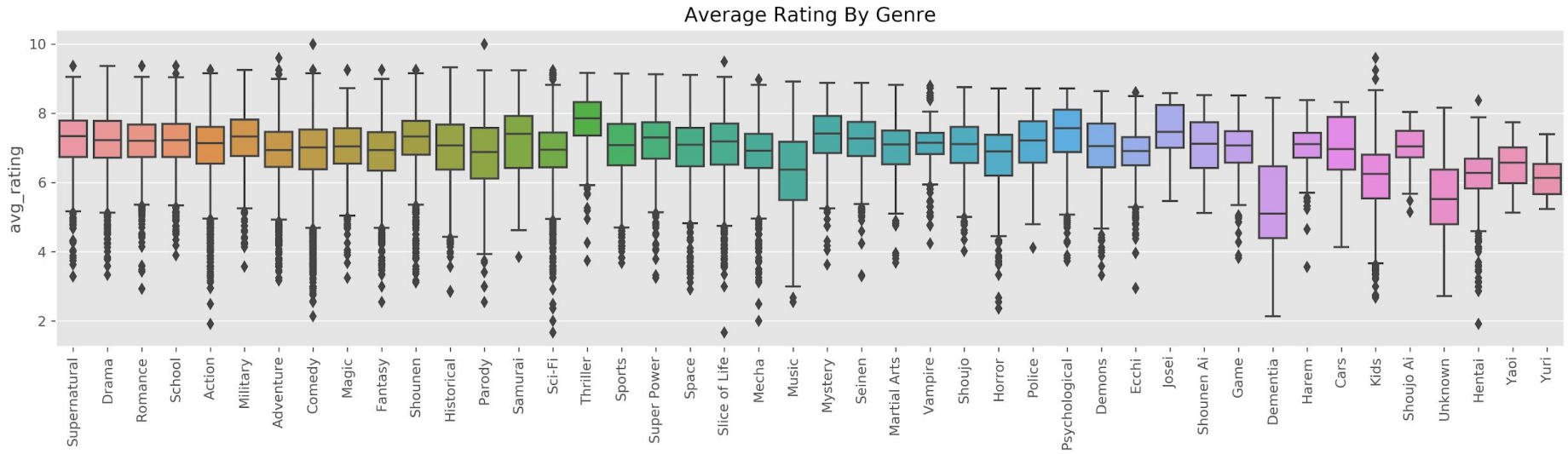
1. Use models to predict expecting rating of a user for an anime, compare with the actual rating provided by the user.
2. Spot-Check some familiar anime, compare results to MyAnimeList recommendations.
3. User Feedback: I will ask a list of anime-enthusiasts to provide feedback on recommendations.

Baseline model using the average rating from data:

- RMSE: 1.57



Average Ratings Across Genre



Some genres perform lower than the majority.



Content Based Recommender System

Iteration 1:

- Features:
Media Type (Movie, TV, etc.), Source (Manga, music, book, etc), Rating Type (PG, R, etc.)
- RMSE (Correlation Similarity Metric): 1.49
- RMSE (Cosine Similarity Metric): 1.35

Iteration 2: Moved forward with cosine similarity based on results from iteration 1

- Features:
Iteration 1 with addition of dummy variables for each genre, and the weighted ratings
- RMSE: 1.31

**Also looked into other iterations, but the above performed the best*

MyAnimeList Recommendations for Hunter x Hunter:



Naruto



Yuu Yuu Hakusho



One Piece

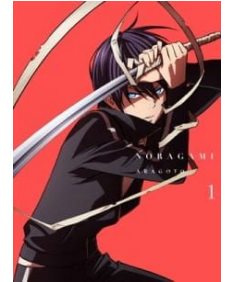
Content Based Recommendations for Hunter x Hunter:



Hunter x Hunter (New)



One Piece



Noragami Aragoto



Collaborative Filter Recommender System

Using Spark's ALS model:

- Final Tuned & Cross Validated Model:
 - Number of latent features: 15
- Results:
 - Train RMSE: 1.03
 - Validation RMSE: 1.15
 - Test RMSE: 1.13

Latent Feature Exploration:

- Latent Features were capturing:
 - Maturity Level
 - Genre
 - Themes: School, Sports, Music

Used Spark's Recommend For All Items feature to create a "Other Users who liked this anime also liked these anime" recommender.



Your Anime Match Maker

[Flask app demo:](#)

Your Anime Match Maker: [Home](#) [Recommender Options](#) ▾

Welcome!

To find your next anime, search for your favorite anime using the form below, or click "Take me to the recommender" if you already know the appropriate anime_id.

[Take me to the recommender!](#)

Need Help Finding the Correct Anime ID?

Select: Movie, TV, or Both:

Search for Title Keyword:

Search



Some Final Notes

User Feedback

- There is a lot of anime
- Anime Enthusiasts: rated recommendations as good.
- For users with less anime knowledge, they rated the recommendations as decent since they did not recognize many of the anime recommended

Caveats:

- The data is only until 2018, missing more recent titles.
- Evaluation of a recommender system is difficult:
 - “Spot-Checking” cases can be subjective
 - RMSE is not a true indicator of whether or not a user will like a recommendation

For the future:

- Add more recent titles.
- NLP on anime descriptions
- User Metadata clustering
 - There was a user dataset, but unfortunately only a small portion of the users in the rating dataset were captured in the user dataset. I would like to look into this further and see if there is another dataset, or if the current one has formatting that is causing the disconnect.
- Neural Network to evaluate RMSE & make recommendations



I hope you find your perfect match.



Sherry Duong



[linkedin.com/in/sherryduong](https://www.linkedin.com/in/sherryduong)



github.com/sherryduong93



animerecz.herokuapp.com



sherryduong@gmail.com

