The Kintoun Anime Recommender System

James Skelton

Goals

- To make an anime recommender using matrix factorization
- Performed analysis using PySpark on AWS cluster
- Share it with friends and the MyAnimeList community

The Source & Data

- MyAnimeList.net is one of the largest, most popular English language review and information websites for Japanese animation.
- The original dataset of user reviews: 31284030 x 11 rows

MyAnimeList

Hide Ads | **Login** | **Sign Up**

Anime   Manga   Community   Industry   Watch   Read   Help

All ▼ | Search Anime, Manga, and more...

One Piece

Details   Videos   Episodes   **Reviews**   Recommendations   Stats   Characters & Staff   News   Forum   Featured   Clubs   Pictures   More Info

Reviews

Write a Review

Sort ⌄

More Reviews

Maxine (All reviews)   Feb 28, 2008
3783 people found this review helpful   344 of ? episodes seen

                                                  Overall Rating: 10

There are two responses I get, without fail, every time I try to get someone new to watch this show. "I don't like the art style" or "I've seen the dub - NO THANKS." I'm guilty of both of these myself. But if there's one thing I need to stress before even getting started on this review, it's that the 4Kids dub is NOT One Piece. For the love of god, PLEASE do not think it is. If you've suffered the misfortune of seeing some of the 4Kids episodes, just erase them from your mind and start fresh. They read more

permalink | report

holdenn (All reviews)   Sep 12, 2008
1460 people found this review helpful   379 of ? episodes seen
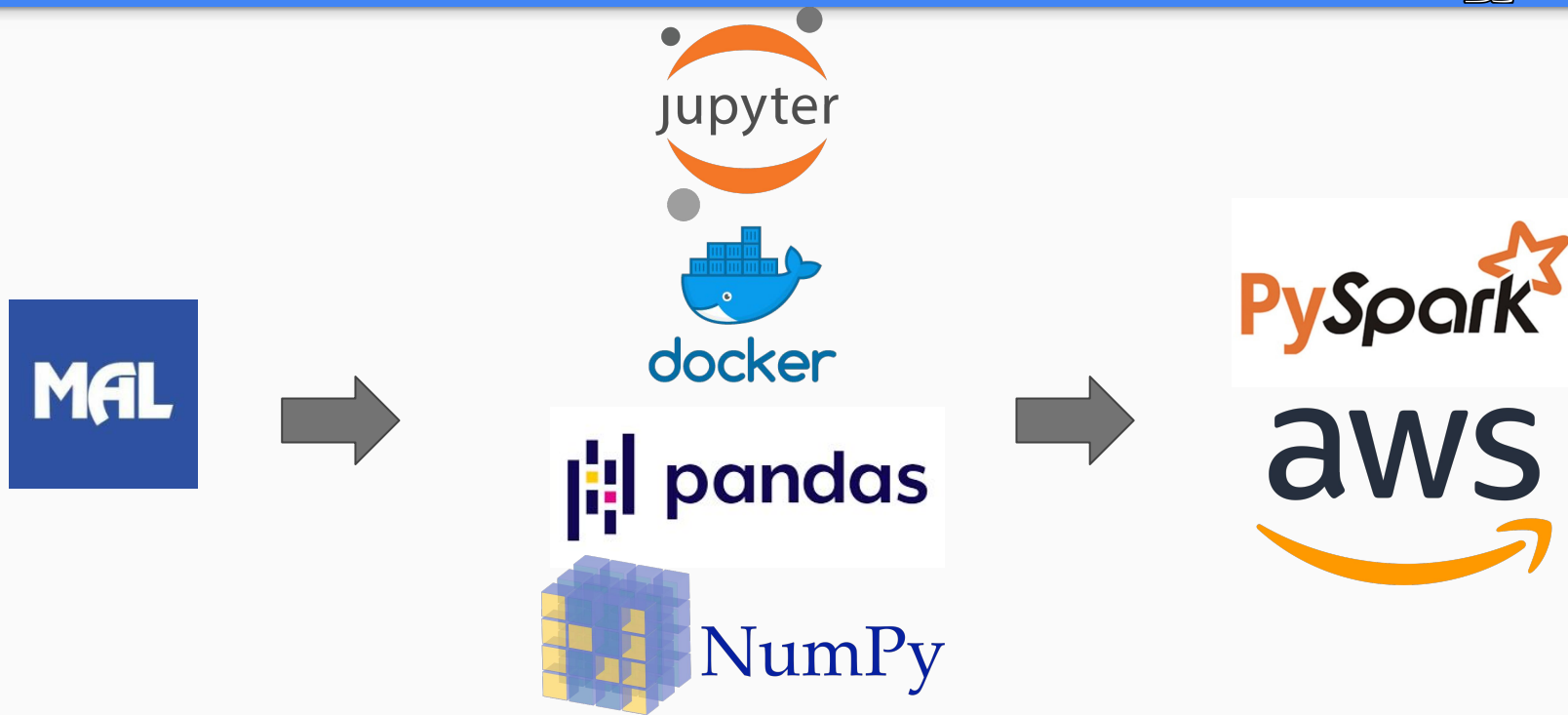
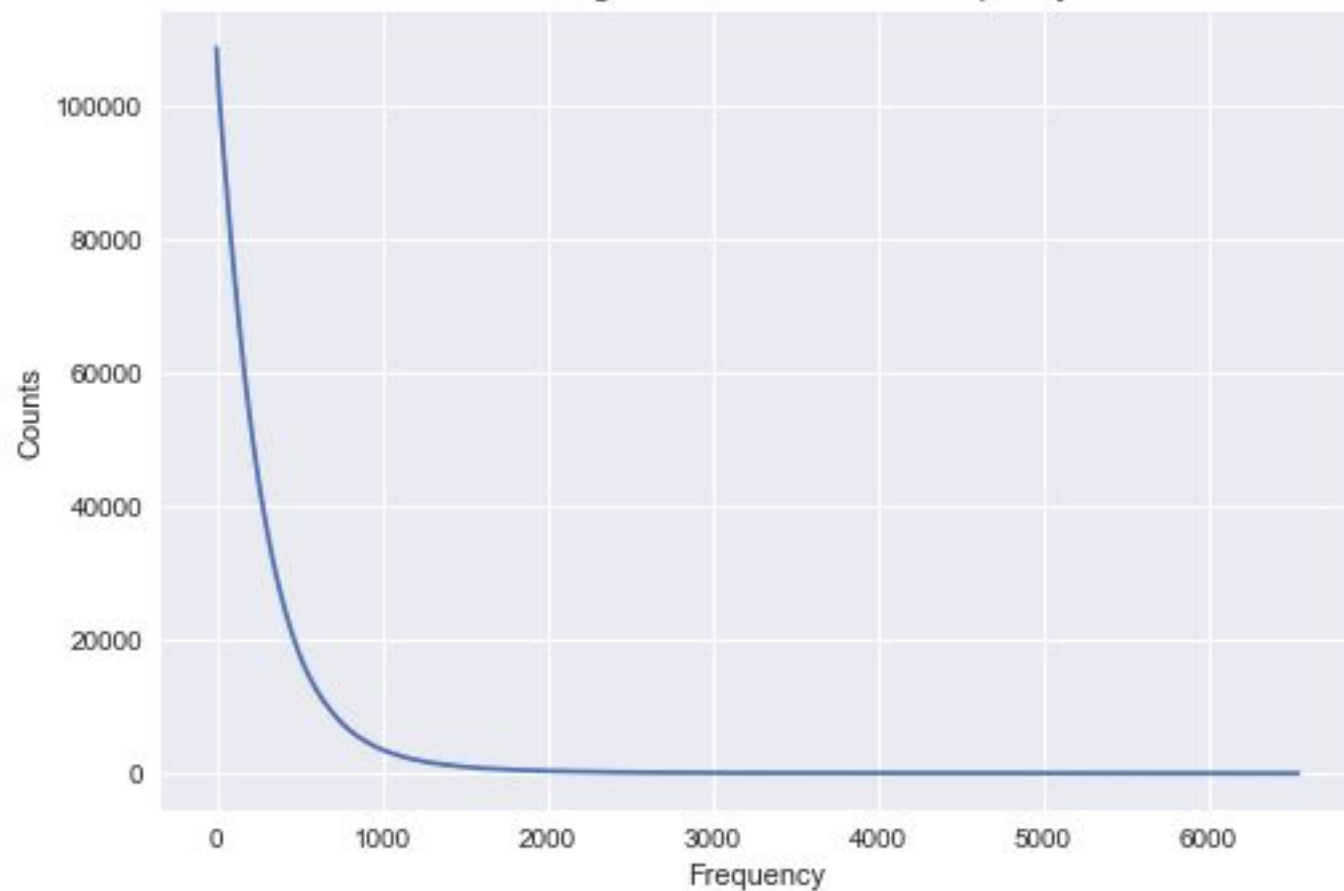   Overall Rating: 10

One Piece is by far the best shounen anime out there that I have watched.
But not all share the same views as I do, lets remedy that, shall we?
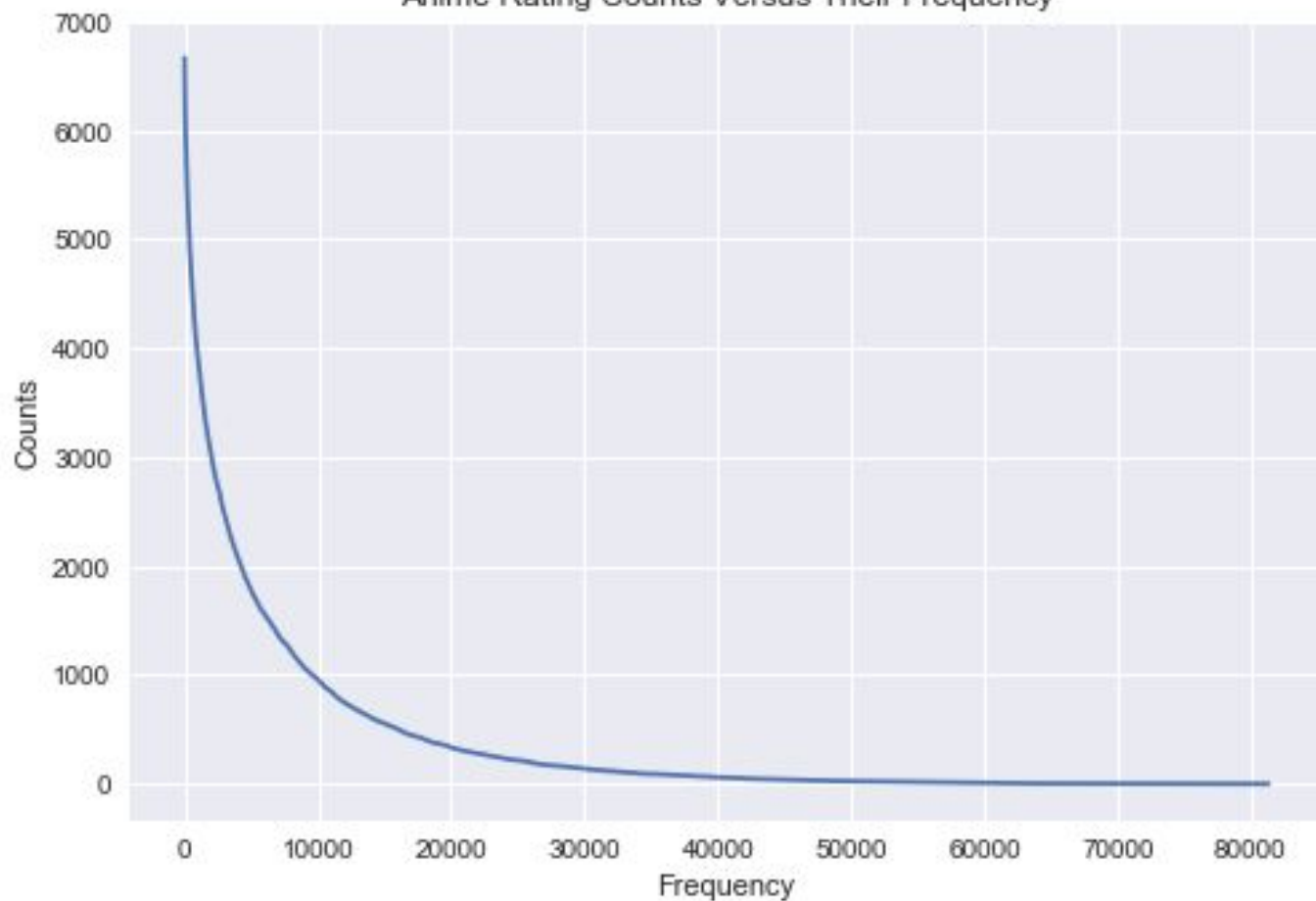
Add to My List

Add to Favorites

Workflow
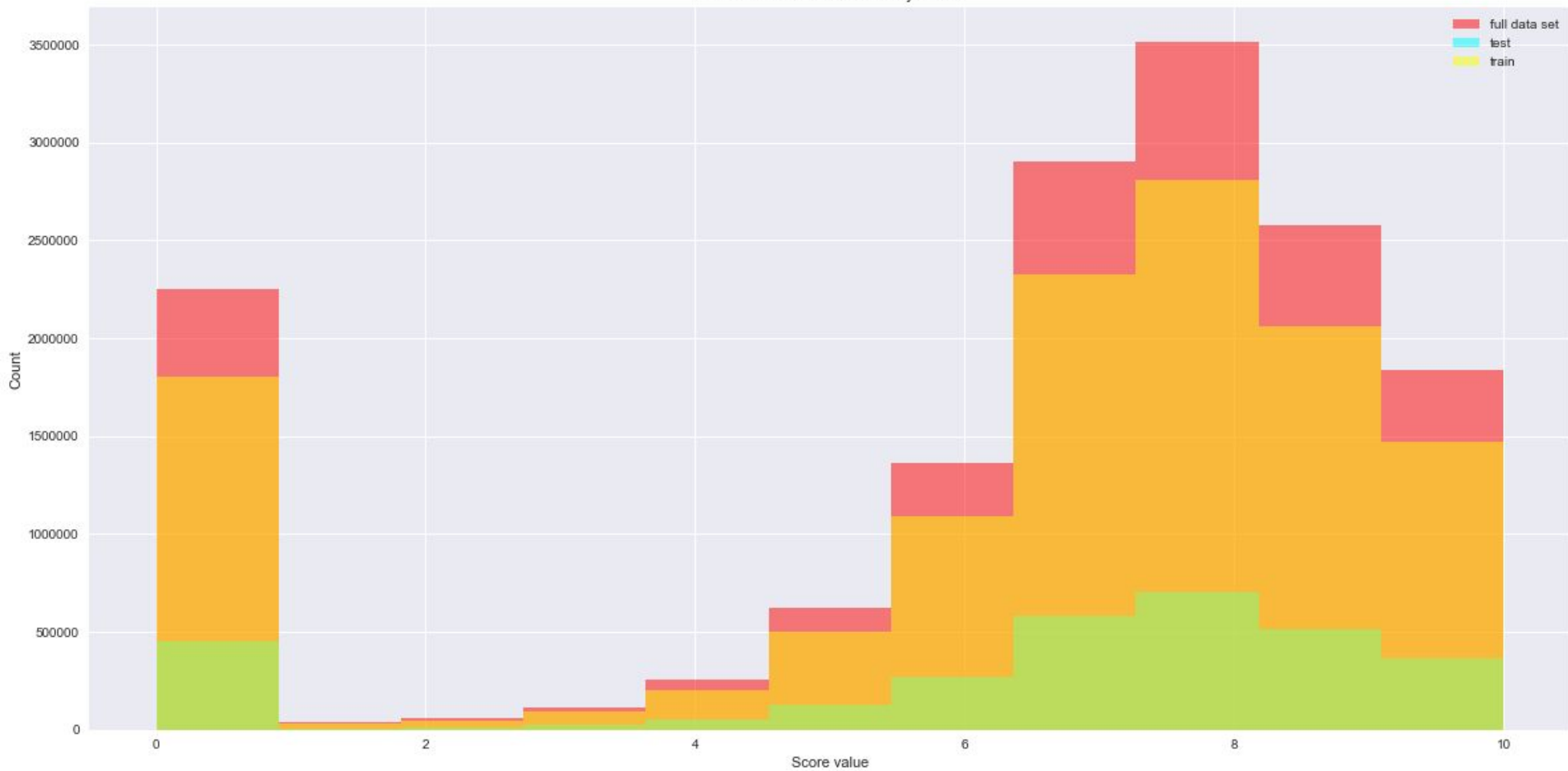
User Rating Counts Versus Their Frequency

Anime Rating Counts Versus Their Frequency
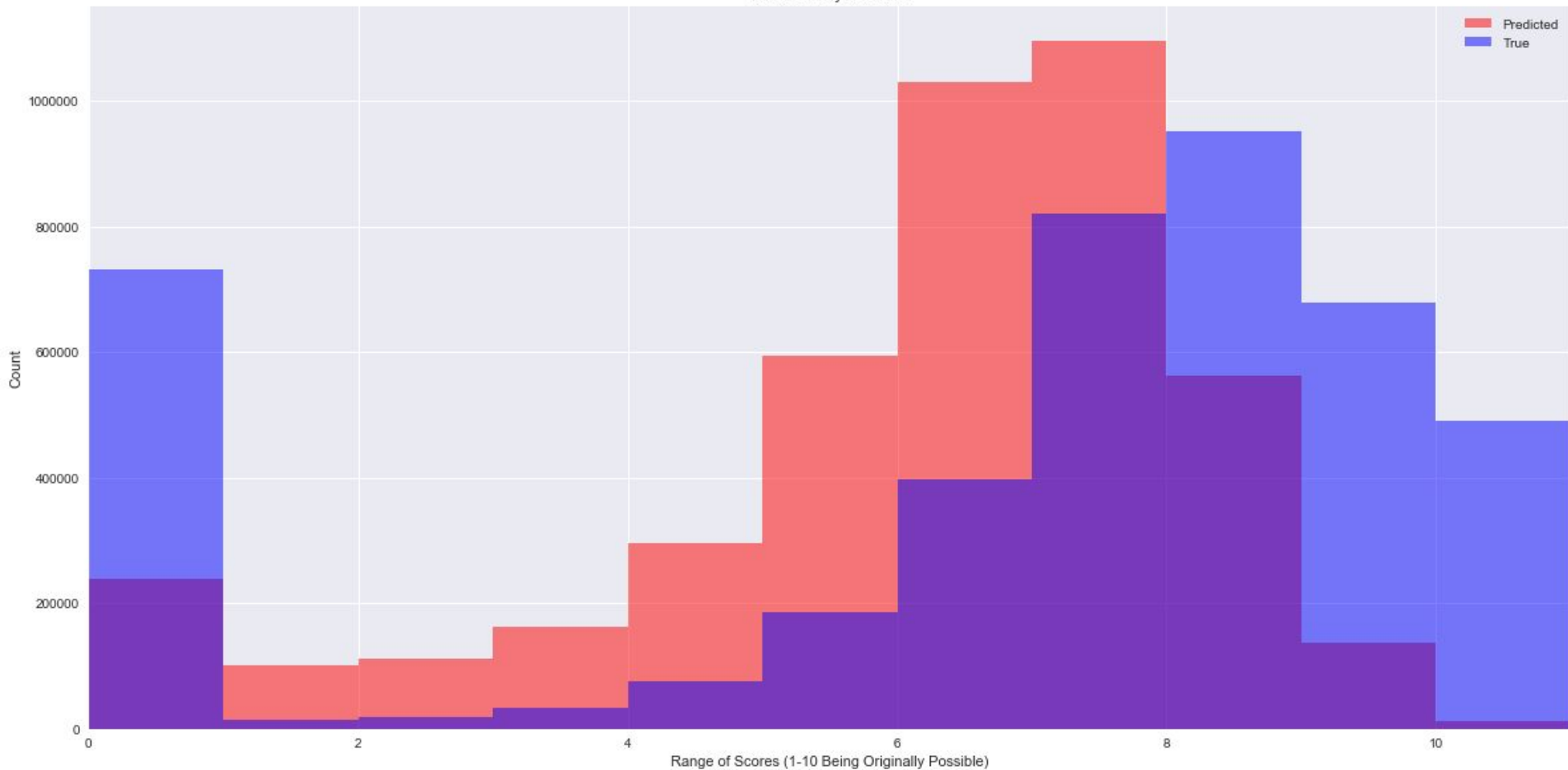
Counts of each Score Given by a User

Key Decisions

- ‘My_status’ > 2 were dropped.
- Only users who’s my_watched_episodes > 3 were retained
- Users with less than 6 reviews were dropped
- Anime with less than 6 reviews were also dropped
- When the recommender was unable to assign a value, the average rating for each anime was used to fill the missing values.
- Final zero count was 16.507% of original

Model & Results

- Spark's Alternating Least Squares matrix factorization model
- The Root Mean Squared Error was 1.614399

Plot of the Predicted versus Actual Scores
Rated on MyAnimeList

Next Steps

- There exists large, unused amounts of information about each anime and user in the original dataset. Incorporating these would offer a number of interesting options to counter further improve this model.
- Another model could be made using this same dataset and Natural Language Processing instead of matrix factorization

Thank you for listening

University of St. Andrews, BSc in Psychology, 2018

Project Github:
https://github.com/Jameshskelton/Anime_Recommender

Linkedin:
https://www.linkedin.com/in/james-skelton-b03aab1a1/

Contact email:
jameshskelton@gmail.com