

Final Project Planning Document

Brian Liles

July 9, 2019

Final Planning Objective

Throughout the DATA 612 course, we have learned several recommender system building techniques. The final project will utilize the collaborative filtering system within a item-based environment.

Data

The data chosen for the final project was downloaded from the Kaggle website

<https://www.kaggle.com/CooperUnion/anime-recommendations-database>

(<https://www.kaggle.com/CooperUnion/anime-recommendations-database>).

```
# rating data set
rating <- read.csv("F:\\CUNY\\Summer 2019\\DATA 612 Recommender Systems\\rating.csv")
glimpse(rating)
```

```
## Observations: 7,813,737
## Variables: 3
## $ user_id <int> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ anime_id <int> 20, 24, 79, 226, 241, 355, 356, 442, 487, 846, 936, 1...
## $ rating <int> -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1, -...
```

The **rating** dataset has 7,813,737 observations and 3 variables.

```
# anime data set
anime <- read.csv("F:\\CUNY\\Summer 2019\\DATA 612 Recommender Systems\\anime.csv")
glimpse(anime)
```

```
## Observations: 12,294
## Variables: 7
## $ anime_id <int> 32281, 5114, 28977, 9253, 9969, 32935, 11061, 820, 15...
## $ name <fct> Kimi no Na wa., Fullmetal Alchemist: Brotherhood, Gin...
## $ genre <fct> Drama, Romance, School, Supernatural, Action, Adventu...
## $ type <fct> Movie, TV, TV, TV, TV, TV, TV, OVA, Movie, TV, TV, Mo...
## $ episodes <fct> 1, 64, 51, 24, 51, 10, 148, 110, 1, 13, 24, 1, 201, 2...
## $ rating <dbl> 9.37, 9.26, 9.25, 9.17, 9.16, 9.15, 9.13, 9.11, 9.10,...
## $ members <int> 200630, 793665, 114262, 673572, 151266, 93351, 425855...
```

The **anime** dataset has 12,294 observations and 7 variables.