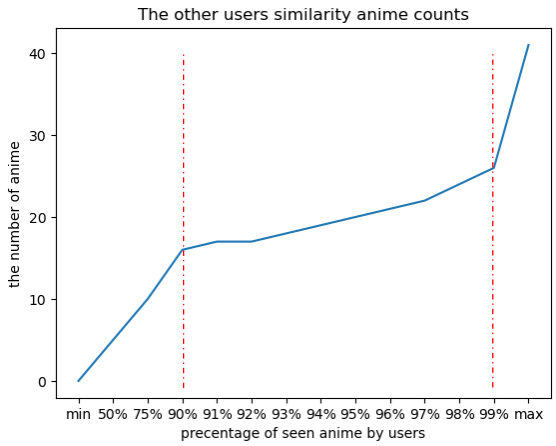
Anime Recommendations Database

Abstract

The recommendation system is one of the most important tasks these days, the company that provides the sale of different products to save their clients (users) needs to have an acceptable recommendation system otherwise, the not satisfied clients will stop to use their service and it means losing revenue. In this project, the author worked on suggesting the most possible anime with collaborative filtering. The author used a simple way to implement it.

Project Description

As checked datasets, the author merged two datasets in the first step and created a pivot table with users as rows and anime as columns and selected one user to work on (randomly selected).



The sample user has given the rate to 41 movies and the distribution of other users has given the rate to similar movies. As you can see the number of movies that are rated by other users has smoothly increased between 90 to 99 percent (from 16 to 26 movies).

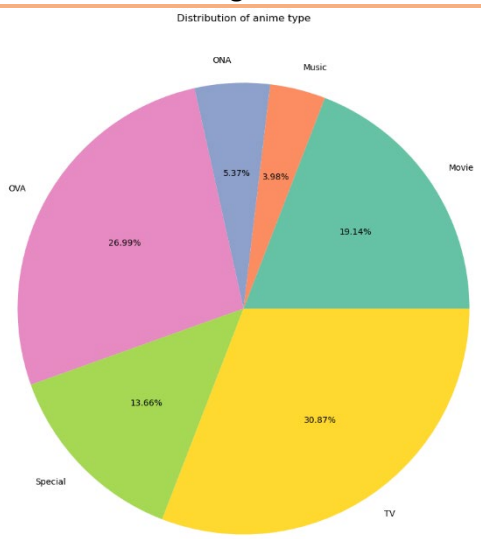
To narrow the similar users, the author used the users that at least rated 25% of the same movie as the target user and found the most correlated rating users with the target user.

To find the list of best movies to recommend, use the rated movies by the most related users and multiply the value of movie rating and users rating and extract the average of rates by sorting it descending to find the list of most possible like movies for the target user.

My exploration

I followed the clustering algorithm on the anime dataset based on some limited features, before that I need to explore and clean my data,

Categorical



Numerical

	episodes	rating	members
count	12064.000000	12064.000000	1.206400e+04
mean	12.253316	6.473902	1.827952e+04
std	46.668767	1.026746	5.527578e+04
min	0.000000	1.670000	1.200000e+01
25%	1.000000	5.880000	2.210000e+02
50%	2.000000	6.570000	1.539000e+03
75%	12.000000	7.180000	9.485500e+03
max	1818.000000	10.000000	1.013917e+06