Part 1

Our project is anime recommender system, focusing on implementing different recommender system based on a anime database dataset.

With the rise of eWorld . The recommender system becomes unavoidable in our daily life. Many famous websites,  such as Amazon, Netflix and Youtube, all have their own unique recommender systems. Accurately recommending the items that users want can generate a huge amount of profit for those companies. Nowadays, there are different kinds of recommender systems, the most famous ones are collaborative filtering and content-based filtering. Collaborative filtering finds similar items based on the user’s  past behaviors, and recommends possible items that the user may like based on history. Content-based recommender is more NLP based and focuses on user’s  profile or item’ description and recommender similar items. There are also other ways of recommending and many industries are using hybrid recommending which combines multiple ways of recommender systems.

 The objective of our project is to try to have experiments with different kinds of recommender systems, compare the performance of some of them if possible.

Below are the main libraries we used for this project, it including some common data analytics library such as sklearn, panda, NumPy , visualization library, matplotliab, seaborn, and recommender system library surprise and also neural network librart keras.

Part 2

For this project, we compare the results of 6 different collaborative filtering recommender system algorithms implemented by suprise, and also including the one recommender net which is implemented by neural network.

. Those 6 models are KNN, KNN with Z score, SVD, SVD plus plus, We use grid search with cross validation to tune parameters and the results are shown as in the bar graph. The evaluation metrics are RMSE and MAE as we need to compare the actual value to the real value and minimize the loss. We find that KNN with Z Score and SVD, SVD plus plus and neural network have the better performance comparing with other algorithms and those 4 algorithms also have similar results.

Besides, we also tried some other methods such as content based, weighted rating. Weighted rating is a method which use a specific formula instead of just using the averaged score to calculate the ranks for all anime. The new formula includes the number of votes for each anime and also the mean scores for all rating which make a more balanced evaluation compared just averaging the score. As we only has the anime information and no user information, we apply content filtering on anime and make item based recommendation. We combine the anime genres and description to make a new scoring and create tf-idf matrix, we calculate the cosine similarities scores based on this matrix and rank all scores and select top one. In the picture, our input is jojo and the output is other jojo series or similar anime. The neural network I just mentioned before just fully connect users and anime and converted into probability after activation function of sigmoid and recommend anime based on the probability.

That is all of our project, thanks.