Recommend System

In this part of the project, I decided to build an item-item collaborative filtering after seeing that there are much more users than games (users:18875, games:2554). By doing item-item filtering, I could save a massive amount of time.

         For building the recommendation system, I primarily use the Surprise library. After load in the dataset, I extracted the necessary columns to form a new data frame. The data frame is later loaded into Surprise’s module: Dataset. I also set the rating scale to (0, 1) through Reader because I am working with a binary dataset. The dataset is then split into a train and test set for encountering overfitting and for evaluation.

         The dataset is fed into five different models using cosine similarity. For KNN families, the models are tuned on two parameters: k and min\_k. The SVD is also tuned on two parameters: n\_epochs and lr\_all. The following is the result:

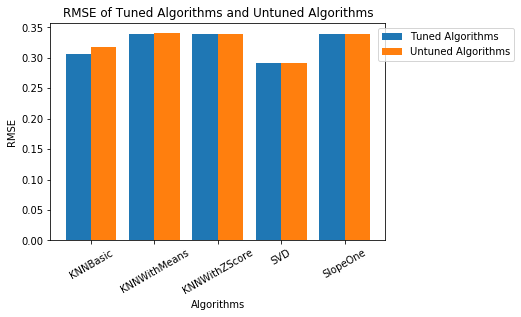


Figure 1. RMSE of Each Models

In conclusion, the SVD has the lowest RMSE thus making the best algorithms to use for building recommender system based on the dataset. We can deploy this algorithm on steam and use A/B testing to see if the new recommender system will significantly outperform the old recommender system. If not, we can try other types of recommender system.