

My hybrid recommendation model is simple, and it mainly contains two parts, one for the item-based collaborative filtering recommendation system and the other for solving the cold start problem.

Specifically, I used the `train_review.json` file to train the item-based model. Notably, this model focuses on co-rated data (i.e., the business pairs need to have at least three co-rated users to qualify for computing the Pearson correlation). Additionally, when making predictions for the user-business pair, I used the nine nearest business neighbors. Also, I used an amplification power to adjust the weight between each pair of `business_id`, which emphasized high weights and punished low weights. This resulting model file are business pairwise similarity.

In the test file, if the `business_id` and/or `user_id` is not included in the model file (i.e., cold start problem), I assigned the grand average of the training dataset as the predicted score.

The last step was tuning the parameters, such as adjusting the proportions of predicted scores contributed by the grand average and item-based model predicted score, to generate the predicted scores that resulted in the least RMSE value.