1. CSE158/258 Cooking Prediction

Kaggle Name : ChiHsin

Method	Description	Validation Result (ACC)	Kaggle Public/Private
Most Popular Item (HW3)	Prediction 1 if item is top percentile% of items train on the 400000 training set. Simply predict 0 if unseen in training set.	percentile : 10, ACC = 0.548565 percentile : 20 , ACC = 0.59391 percentile : 30 , ACC = 0.634985 percentile : 40 , ACC = 0.66807 percentile : 50 , ACC = 0.69627 percentile : 60 , ACC = 0.723925 percentile : 70 , ACC = 0.754945 percentile : 80 , ACC = 0.783615 percentile : 90 , ACC = 0.8065	percentile : 50 , 0.68660/0.68680 percentile : 60 , 0.69930/0.70190 percentile : 70 , 0.70230/0.70540 percentile : 80 , 0.68480/0.68810
Jaccard (HW3)	Compute the Jaccard Similarity of all items cooked by user in training set wrt the item to predict, and if the maximum similarity exceeds threshold, predict 1. Predict 0 othrewise.	threshold : 10 , ACC = 0.8077 threshold : 20 , ACC = 0.75093 threshold : 30 , ACC = 0.70707 threshold : 40 , ACC = 0.655665 threshold : 50 , ACC = 0.595455 threshold : 60 , ACC = 0.59426 threshold : 70 , ACC = 0.58028 threshold : 80 , ACC = 0.57944 threshold : 90 , ACC = 0.57944	threshold : 10 , 0.50849/0.50310 threshold : 20 , 0.52500/0.51920

2. CSE258 Recipe Rating Prediction

Kaggle Name : ChiHsin

Method	Description	Validation Result (MSE)	Kaggle Public/Private
Bias only Latent Factor Based Model (Workbook 4)	As descried in hw3 problem 11. Compute mean, user and item biases on training set, and run gradient descent with regularization.	lamb = 1e−05, MSE = 0.6343308205994943 lamb = 0.0001, MSE = 0.7800174629415336 lamb = 0.001, MSE = 0.8574004686509782 lamb = 0.01, MSE = 0.8911307388311662 lamb = 0.1, MSE = 0.8996097642865012 lamb = 1, MSE = 0.9007586515538737 lamb = 10, MSE = 0.9008789032254846 lamb = 100, MSE = 0.9008909863002668 lamb = 1000, MSE = 0.900892195231244	lamb = 1e−5, MSE = Turned out I generated the wrong file for this. Possibly I could beat the strong baseline by more with this one if I had generated the right one? IDK. lamb = 1, MSE = 0.91834/ 0.87023
Latent Factor Model	Using Surprise SVD model.	MSE = 0.822901845396487	0.83852/0.81665
Baseline	Predict user mean if user seen in training set, otherwise predict global average.	MSE = 0.7689307678860396	0.83256/0.81882
Mean of Baseline and Latent Factor Model	Predict score that is mean of above two methods.	MSE = 0.8068564418316578	0.84387/0.81884