Project 2

Angeline Chen

	MAE of GIVEN 5	MAE of GIVEN 10	MAE of GIVEN 20	OVERALL MAE
User Cosine Similarity	0.857821683131174	0.7925	0.769460789042153	0.804137251682811
User Pearson Correlation	0.875328248093035	0.816833333333333	0.774573164850005	0.818051223116073
User Pearson Correlation IUF	0.872327122671002	0.815666666666667	0.806212018906144	0.830241339681497
User Pearson Correlation Case	0.879829936226085	0.834333333333333	0.806212018906144	0.837300935806928
Item Cosine Similarity	0.876078529448543	0.802833333333333	0.801099643098293	0.826136923329503
Ensemble	0.790421408028011	0.754333333333333	0.733384778624481	0.757264816943031
Deep Learning	0.92872327122671	0.935333333333333	0.902671939809009	0.919266130356263

I think that my results are reasonable because the MAEs of the methods are similar.

Cosine similarity and Pearson correlation measure the similarity between users. Pearson correlation improves on cosine similarity because it centers the data by subtracting the average of the user's ratings. IUF improves on Pearson correlation because it penalizes universally liked movies. Case amplification improves on Pearson correlation because it favors high weights. However, the improvements do not necessarily decrease the MAE because the performance depends on other parameters, such as the number of similar users. I defined similar users as all train users who rated at least one movie that the test user rated.

User-based collaborative filtering assumes that similar users rate similarly. Item-based collaborative filtering assumes that similar items are rated similarly. Because the MAEs of user-based collaborative filtering is less than the MAEs of item-based collaborative filtering, the assumption that similar users rate similarly is better than the assumption that similar items are rated similarly for movie predictions.

I implemented an ensemble algorithm and a deep learning algorithm. The ensemble algorithm uses the average rating of the methods with the least MAE, cosine similarity and improved Pearson correlation, to predict the rating. Smoothing improves on Pearson correlation. If a user rated more movies, it assigns higher weights to the average of the user's ratings. If a user rated fewer movies, it assigns lower weights to the average of the user's ratings. The MAE of the ensemble algorithm is less than the MAE of the methods used. The deep learning algorithm optimizes the weights by training the neural network. In the input layer, the inputs of the neural network are the train user's ratings. In the hidden layers, the neurons assign weights to the inputs. The weights are initialized to the cosine similarity and are adjusted to minimize the cost function during training. The cost function is the difference between the train user's rating and the output of the sigmoid function. In the output layer, the output of the neural network is the sum of the weights. The MAE of the deep learning algorithm is greater than the MAE of the other methods because the weights are adjusted to different values each time. When I used cross validation to evaluate the performance of the deep learning algorithm, the range of the MAEs of the deep learning algorithm was greater than the range of the MAEs of the other methods.