Assignment 1 Due Date – 14th February 12 Noon 10 marks

Download the Movielens dataset 1 (the 100K dataset) from http://www.grouplens.org/node/73

Build a simple user-based and item-based recommender system (papers have been provided).

The dataset will have available and missing ratings. For evaluation you should use 5 fold cross validation. In each run, use 4 parts for training and the remaining 1 part for testing. Use the training set to predict the ratings of the test set.

For user based approach (table 1) see how the MAE changes when the threshold for neighborhood selection varies: use thresholds 0.4, 0.5, 0.6, 0.7

For the item based approach (table 2) see how the accuracy changes when the number of similar items are restricted to K most similar items. Vary K: 10, 20, 30, 40

To test how good or bad your recommender system is, you should compute the Mean Absolute Error (MAE) on the test set.

For measuring similarity you should use the **Cosine Similarity**

Table 1. MAE values User based

Fold #	τ=.4	τ=.5	τ=.6	τ=.7
1	0.827554174644 26	0.827530939003 3723	0.827542337859 6213	0.827570468667 0641
2	0.820188056557 3212	0.820183932260 4757	0.820177840826 9204	0.820255913039 0149
3	0.810859395979 4573	0.810853704967 1175	0.810885740973 8569	0.810863900655 9971
4	0.809576884347 7015	0.809573098939 4443	0.809489101897 2152	0.809427678408 6285
5	0.814246428290 9688	0.814246548795 1474	0.814230576937 8908	0.814212548092 346
Average	0.81648498796	0.81647764479	0.81646511969	0.81646610177

Table 2. MAE values Item based

Fold #	K=10	K=20	K=30	K=40
1	1.079672638888 8884	1.097047657171 2606	1.109273459552 6967	1.116515810481 5145
2	1.065570337301 587	1.088287279930 037	1.102632964809 8856	1.108654260231 7127
3	1.040049047619 0476	1.054382447217 8156	1.061594193597 5013	1.066844334086 7804
4	1.034474404761 9054	1.050565914604 4022	1.058537062081 2244	1.063964437894 584
5	1.033865019841 2693	1.051641639198 2208	1.059162572225 5642	1.065102116282 0047

Average	1.05072628968	1.06838498762	1.07824005045	1.0842161918

- 5+5 marks