

Download the Movielens dataset 1 (the 100K dataset) from:

<https://grouplens.org/datasets/movielens/100k/>

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.

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**

Report how the MAE changes when the number of neighbours 'K' varies (K = 10, 20, 30, 40, 50) for each both user based and item based approaches.

You have to fill two tables (one for user based and one for item based) of the following form

	K=10	K=20	K=30	K=40	K=50
Fold 1					
Fold 2					
Fold 3					
Fold 4					
Fold 5					
Average					

You have to submit a PDF file. You will also have to demonstrate the code in front of TA.

- 10 marks

Bonus

If you implement significance weighting and variance weighting and show the results, you will receive one bonus mark for each type, i.e. if you implement both you will get 2.