

Feedback — Programming Assignment 4

[Help](#)

You submitted this homework on **Fri 22 Nov 2013 12:28 AM PST**. You got a score of **12.00** out of **12.00**. However, you will not get credit for it, since it was submitted past the deadline.

This quiz is the questions for [Programming Assignment 4](#).

Question 1

Why do PersMean and ItemMean have the same nDCG?

Your Answer	Score	Explanation
<input checked="" type="radio"/> Because the user mean does not affect ranking for a particular user.	✓ 1.00	
<input type="radio"/> No reason; it is an accident of this data set		
<input type="radio"/> Because personalized mean is not a good algorithm.		
Total	1.00 / 1.00	

Question Explanation

nDCG measures rank accuracy. Since the user mean does not change from item to item, only the item mean is used to rank items in the personalized mean, so they will rank items identically.

Question 2

How many neighbors are needed for normalized Lucene to beat personalized mean on RMSE?

You entered:

Your Answer	Score	Explanation
50	✓ 1.00	

Total 1.00 / 1.00

Question Explanation

We awarded points for values between 41 and 51.

Question 3

How many neighbors are needed for normalized Lucene to beat personalized mean on prediction nDCG?

You entered:

75

Your Answer		Score	Explanation
75	✓	1.00	
Total		1.00 / 1.00	

Question Explanation

We awarded points for values between 67 and 77.

Question 4

When is Lucene CBF the best algorithm?

Your Answer		Score	Explanation
<input type="radio"/> Never			
<input type="radio"/> On RMSE with a medium number of neighbors			
<input checked="" type="radio"/> On nDCG, when it is normalized and has many neighbors	✓	1.00	
Total		1.00 / 1.00	

Question 5

What algorithm produces the most diverse Top-10 lists (by our entropy metric)?

Your Answer	Score	Explanation
<input type="radio"/> Popular Movies		
<input type="radio"/> User-User CF		
<input type="radio"/> Normalized Lucene		
<input type="radio"/> Unnormalized Lucene		
<input checked="" type="radio"/> Item/Pers. Mean	✓ 1.00	
Total	1.00 / 1.00	

Question 6

Does increasing neighborhood size generally increase or decrease the tag entropy of top-10 lists?

Your Answer	Score	Explanation
<input checked="" type="radio"/> Increase	✓ 1.00	
<input type="radio"/> Decrease		
Total	1.00 / 1.00	

Question 7

Why is Top-N nDCG lower than prediction nDCG?

Your Answer	Score	Explanation
<input type="radio"/> Because the rating predictor, tested by prediction nDCG, clamps predictions to be in the 0.5-5 star rating range		
<input checked="" type="radio"/> Because the recommender may put items the user has not rated at the front of the list	✓ 1.00	
<input type="radio"/> Because the top-N nDCG is a more realistic test		
<input type="radio"/> Because it is considering 10 items rather than 5		

Total	1.00 /
	1.00

Question Explanation

Top-N nDCG counts unrated items, along with low-rated items, as 'irrelevant'. Prediction nDCG only considers rated items, and measures how well the recommender does at ranking them. For this reason, top-N nDCG is lower, as it may push additional items above the 'good' items in the ranking.

Question 8

In practice, recommenders cost money to run and it isn't worthwhile to run recommenders that take a lot of computational power and provide little benefit. Based on this experiment, what algorithm would be best to deploy for recommending items (in ranked lists) from this data set?

Your Answer	Score	Explanation
<input type="radio"/> Normalized Lucene		
<input type="radio"/> Unnormalized Lucene		
<input checked="" type="radio"/> Popular	✓ 1.00	
<input type="radio"/> Mean Rating		
<input type="radio"/> User-user CF		
Total	1.00 / 1.00	

Question Explanation

The purpose of this question was to have you think about what you have learned in this module and interpret the results of your evaluation. The improvement of the best personalized algorithm vs. Popular was very small, and Popular is a very efficient algorithm, so the available data suggest Popular is the best algorithm for generating top-N recommendation lists on this data. If you want measurements of time, the key one to look at in this case is the TestTime. It measures how much time the algorithm takes to generate all the predictions in a particular test run. BuildTime measures how long it took to compute the model, but is not very relevant for most of the recommenders we considered.

Question 9

Ignoring entropy, what user-user configuration generally performs the best?

Your Answer	Score	Explanation
<input type="radio"/> Normalized		
<input type="radio"/> Unnormalized		
<input checked="" type="radio"/> Normalized w/ Cosine	✓ 1.00	
Total	1.00 / 1.00	

Question Explanation

Normalized user-user with cosine similarity performed the best for a wide range of neighborhood sizes on most metrics.

Question 10

One algorithm has increasing entropy for low neighborhood sizes, and then entropy starts going down. Which is it?

Your Answer	Score	Explanation
<input type="radio"/> Unnormalized User-User CF		
<input type="radio"/> Unnormalized Lucene		
<input checked="" type="radio"/> Lucene with Normalization	✓ 1.00	
<input type="radio"/> User-user CF with Cosine		
Total	1.00 / 1.00	

Question 11

What algorithms beat Popular Items in recommendation list entropy?

Your Answer	Score	Explanation
<input type="checkbox"/> Unnormalized Lucene	✓ 0.25	
<input checked="" type="checkbox"/> Normalized Lucene	✓ 0.25	
<input checked="" type="checkbox"/> User-user CF	✓ 0.25	
<input checked="" type="checkbox"/> Item/Pers. Mean	✓ 0.25	

Total

1.00 / 1.00

Question 12

What is the Top-N nDCG of Popular?

You entered:

0.474

Your Answer		Score	Explanation
0.474	✓	1.00	
Total		1.00 / 1.00	

Question Explanation

We awarded points for values between 0.45 and 0.50.