Congratulations! You passed!

TO PASS 80% or higher

Keep Learning

GRADE
90.90%

Offline Evaluation and Metrics Quiz

LATEST SUBMISSION GRADE
90.9%

1. Why would you use a different metric for evaluating prediction vs. top-N recommendation? **1 / 1 point**

✓ **Correct**

2. Which of the following is an advantage of nDCG *compared with Spearman rank correlation*? **1 / 1 point**

✓ **Correct**

3. Which of the following statements about diversity metrics is **not true**? **1 / 1 point**

✓ **Correct**

4. In some top-n evaluations, instead of considering all items, the recommender recommends from the items the user has rated/consumed/purchased plus a random subset of all items. Why is this useful? **1 / 1 point**

✓ **Correct**

5. Which of the following is a true statement about why someone might prefer to use RMSE (Root Mean Squared Error) instead of MSE (Mean Squared Error) or MAE (Mean Absolute Error)? **0 / 1 point**

! **Incorrect**

6. When computing serendipity, we depend upon a prior “primitive” estimate of obviousness and a determination of whether a recommended item is actually relevant. Why do we need these measures? **1 / 1 point**

✓ **Correct**

7. What is the **major** problem of offline evaluation with unary data? **1 / 1 point**

✓ **Correct**

8. When holding out ratings from a user’s profile for evaluation, what is the benefit of holding out the last ratings rather than holding out random ratings? **1 / 1 point**

✓ **Correct**

9. You've learned about many techniques for evaluation. We also pointed out that most evaluation techniques do not address the question of whether the items recommended are actually useful recommendations. Instead, those evaluations focus on whether the recommender is successful at retrieving “covered up” old ratings. Which of the following evaluation metrics successfully focuses on whether the recommender can produce recommendations for new items that haven’t already been experienced by the user? **1 / 1 point**

✓ **Correct**

10. What is the purpose of decision-support metrics such as reversals, precision, or ROC? **1 / 1 point**

✓ **Correct**

11. Which of these statements best explains how we perform an n-fold cross validation for getting a more accurate measure of the accuracy experienced by users in a recommender system? **1 / 1 point**

✓ **Correct**