Midterm Part 3

12 questions

1 point

1. Why is average rating often inappropriate for ranking items in a non-personalized recommender?

- a. Because the item may have no ratings.
- b. Because a small number of ratings is not enough evidence that the item is particularly good.
- c. Because popular items should be ranked first.
- d. Because the crowd has bad taste.

1 point

2. Resnick discussed a sybil-based shilling attack against a recommender system. Which of these best describes such an attack?

- a. Rating items randomly to confuse the recommender
- b. Creating bogus accounts to promote (or demote) particular items
- c. Creating many accounts to overload the recommender
- d. Writing a review of a book you wrote with your personal account, but hiding your identity

1 point

3. Which of the following recommender algorithms uses item attributes such as movie genres or actors?

- a. User-user collaborative filtering
- b. Non-personalized summary statistics
- c. Product association recommenders
- d. Content-based filtering

---

1 point

4. The reading “How Not to Sort by Average Rating” argues against the way many websites rank rated items. What does the article argue is the best measure to use for such ranking?

- a. The total number of positive ratings, ignoring negative ratings.
- b. The lower bound of a confidence interval around an estimated real fraction of positive ratings.
- c. The percentage of ratings that are positive (i.e., positive ratings / total ratings).
- d. The net number of positive ratings (i.e., positive ratings - negative ratings).

---

1 point

5. We claimed that some Zagat fans feel the guide is getting worse due to self-selection bias and greater diversity of users. What does this mean?

- a. Many restaurants have too many ratings to be able to compute an average score.
- b. Restaurants that are “not that good” get high scores because they are only rated by the subset of people who enjoy them.
- c. People mostly choose to go to highly-rated restaurants, hurting the ability to compute scores for other restaurants.
- d. People have very different ratings distributions, which messes up the computation of average scores.

1

6. Which of these statements best describes the difference between predictions and recommendations?

point

- a. Recommendations focus on products you're likely to consume; predictions focus on matching your past ratings.
- b. Recommendations are based on implicit ratings; predictions are based on explicit ratings.
- c. Recommendations are selected items (or lists of items); predictions supply scores for any particular item.
- d. Recommendations are the result of using a recommender system; predictions come from machine learning models.

---

1 point

7. What is meant by an “organic” prediction or recommendation?

- a. It applies to products that have been pre-filtered (derived from the grocery industry which used the term for recommenders of organic produce and meats).
- b. It applies to recommendations based only on ratings from other users, without including any outside data (such as box-office sales) or attribute data.
- c. It applies to the display of a set of products together, either in a list or a two-dimensional layout display.
- d. It applies to the mode of displaying the recommendation or prediction--organic displays fit naturally into the product display interface rather than announcing “this is a recommendation” or “here are some predictions for you”.

---

1 point

8. When is “term-frequency” most useful as part of a content-filtering recommender?

- a. When certain items are much more popular than other items.
- b. When users are unlikely to have experienced many of the items in the system.
- c. When the attributes of the items can apply in different degrees to different items.
- d. When certain terms aren't very useful because they apply to too many different items.

1 point

9. User-user collaborative filtering depends on certain assumptions. Which of the following is NOT a requirement for a successful user-user collaborative filtering system.

- a. User tastes must either be generally stable (individually) or if changing, they change in sync with other user's tastes.
- b. The domain in which we are performing collaborative filtering is scoped such that people who agree within one part of that domain generally agree within other parts of the domain.
- c. Users mostly have similar tastes on a set of popular items, though they may have individually different tastes on unpopular items.
- d. Past agreement between users is predictive of future agreement -- i.e., if you and I have agreed on items before, we mostly still do now.

1 point

10. A more advanced user-user collaborative filtering formula is:

$$\mathrm{P}_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{n} (r_{u,i} - \bar{r}_u) \times \mathrm{w}_{a,u}}{\sum_{u=1}^{n} \mathrm{w}_{a,u}}$$

What is the purpose of the $\bar{r}_a$ and $\bar{r}_u$ terms in this version of the formula?

- a. These terms specify that we're combining the ratings of lots of other users together.
- b. These terms weight the recommendations so closer neighbors count more than distant neighbors.
- c. These terms limit the number of neighbors used in the computation.
- d. These terms normalize the computation to adjust for different users' rating scales.

1 point

11. Consider the idea of enhancing the Zagat restaurant guide with personalized recommendations. Which of the following statements is true about making that enhancement?

- a. We couldn't do content-based filtering without exposing the identities of each of the restaurant raters.
- b. We couldn't do user-user collaborative filtering without a thorough set of restaurant attribute data such as cuisine, price level, etc.
- c. We couldn't do user-user collaborative filtering without being able to associate ratings with individual users or user IDs.
- d. We can do personalized recommendations entirely by using summary statistics such as mean and standard deviation of ratings.

1 point

12. Golbeck explained that trust-based recommenders differ from similarity-based collaborative filtering in all of these way EXCEPT which one?

- a. Trust-based systems are harder to get going, because it is often challenging to get trust data.
- b. Similarity-based collaborative filtering treats all rated items as roughly equivalent in evaluating neighbors, trust-based systems may give very strong weight to the items that a user is most passionate about.
- c. Trust-based systems only consider ratings from users that the target user has a direct trust relationship with, and thus often use many fewer ratings in computing a prediction or recommendation.
- d. Trust-based systems have an underlying graph of user trust, while similarity-based systems don't need a graph because they only use pairwise similarity scores.

Submit Quiz