Midterm Part 1

12 questions

1 point

1. We've used the example of recommending bananas in a grocery store as an example of how not to build an effective recommender. What is wrong with recommending bananas?

- a. Very few people actually like bananas, so they're unlikely to trust the recommender.
- b. Bananas and other fruit are perishable; you generally only want to recommend things that can stay in inventory for a long time.
- c. People usually buy bananas anyway without a recommendation, so why waste the opportunity to recommend something new.
- d. Bananas are usually too inexpensive to be worth recommending -- try recommending something expensive like lobster or steak instead.

1 point

2. A user views the first 13 seconds of a 5 minute video on YouTube, then browses away. What kind of recommendation input is this?

- a. Implicit feedback
- b. Unary data
- c. A rating
- d. Indication that the user likes the movie

1 point

3. Which of the following is a problem with using Pearson correlation (as opposed to other similarity metrics) for computing user similarities in user-user collaborative filtering?

- a. Users may use different portions of the rating scale.
- b. Users may not have rated any of the the same items.
- c. If users have only rated a small number of the same items, their correlation may be too high.
- d. The user may not know any other users in the system.

1 point

4. Which of the following is NOT true about case-based reasoning?

- a. It searches a base of prior knowledge to find items the user is likely to find interesting
- b. It provides a natural basis for interactive, conversational recommendation where the user iteratively refines their browsing
- c. It infers item similarity from user ratings
- d. It characterizes items by features

1 point

5. Conde Nast Traveller computes non-personalized scores as the percentage of people who rate a particular item “very good” or “excellent.” In what way is this BETTER THAN Zagat’s average rating model?

- a. It uses real ratings from users.
- b. It avoids giving too much weight to extreme ratings.
- c. It is easy to explain.
- d. It links recommender scores to product attributes.

1 point

6. We provided two product association recommender formulas: A simple formula $\frac{(\text{X and Y})}{\text{X}}$ and an adjusted one $\frac{\left(\frac{(\text{X and Y})}{\text{X}}\right)}{\left(\frac{(!\text{X and Y})}{!\text{X}}\right)}$. In which circumstance might we prefer the simple formula?

- a. When the products being recommended have been purchased, not just rated.
- b. When users are familiar with most popular products and are looking for new, less-popular ones.
- c. When the X and Y scores are derived from ratings, not clicks or purchases.
- d. When it seems reasonable to recommend popular items as long as they are associated with the target item.

1 point

7. Which of these techniques is NOT used for building a content filtering profile?

- a. Build an attribute preference vector based on the most popular items in the catalog.
- b. Build an attribute preference vector from explicit user ratings.
- c. Build an attribute preference vector based on user actions such as view/click/buy.
- d. Provide an interface where users can specify and edit their own vector.

1 point

8. Which of these statements best describes the goal of the TFIDF formula?

- a. To select items that most match your specific preferences, weighting each preference more highly if only a small number of items match that preference.
- b. To select items that most match your specific preferences, weighting each preference more highly if a large number of items match that preference.
- c. To select items that most match your specific preferences, weighting each preference equally.

- d. To select items for you that generally match your preferences, but weighting the selection to favor items that are not selected by users very often.

1 point

9. Either vector cosine or Pearson correlation are often used to compute a weight in user-user collaborative filtering. What are these metrics trying to measure?

- a. These are measures of the similarity of ratings history between users.
- b. These are measures of the number of ratings users have in common.
- c. These are measures of how much the target user likes popular items.
- d. These are measures of how well the recommendations match the user's preferences.

1 point

10. Which of these explanations of a user-user collaborative filtering predicted rating is most likely to be effective?

- a. A detailed explanation of the prediction algorithm, including an explanation of Pearson correlation and normalization of ratings.
- b. A simple chart showing how many users like, disliked, and were neutral about the item.
- c. A scatterplot showing other users' ratings of the item on one axis, and their correlation with the target user on the other axis.
- d. A detailed look at the closest neighbor, including that neighbor's correlation with the target user and that neighbor's rating of the item.

1 point

11. A basic user-user collaborative filtering algorithm uses the formula:

$$\mathrm{P}_{a,i} = \frac{\sum_{u=1}^{n} r_{u,i} \cdot \mathrm{w}_{a,u}}{\sum_{u=1}^{n} \mathrm{w}_{a,u}}$$

What is the purpose of the term $w_{a,u}$ in the numerator?

- a. It is used to normalize ratings, since users rate on different scales.
- b. It is used to make sure that only a limited number of neighbors are part of the computation.
- c. It is used to give some neighbors a greater influence on a target user's prediction than others.
- d. It is used to ensure that the resulting prediction is on the same scale as the ratings.

---

1 point

12. Cosley experimented with giving people deliberately inaccurate predictions. He examined three possibilities:

I. People would notice that predictions were wrong

II. People would be biased by the wrong predictions and enter different ratings.

III. People would have lower satisfaction with the system after receiving bad predictions.

Which ones happened?

- a. All three of the results were confirmed.
- b. I and II were confirmed, but III was not confirmed.
- c. I and III were confirmed, but II was not confirmed.
- d. II and III were confirmed, but I was not confirmed

Submit Quiz