



## Midterm Part 2

12 questions

1  
point

1. Which of these statements about information filtering and information retrieval is NOT TRUE?

- ☐ a. Information filtering focuses on building profiles of long-term user interest while information retrieval focuses on building indexes of content.
- ☒ b. Information filtering involves matching a user-entered query to document terms or item attributes.
- ☐ c. Information filtering evaluates new content items for match against user profiles.
- ☐ d. Information retrieval often uses the TFIDF approach where terms are more relevant if they occur in few documents, but frequently in the matched document.

1  
point

2. In our taxonomy of recommender systems, what do we mean by “ephemeral personalization?”

- ☐ a. Ephemeral personalization means that you may get different recommendations the next time you log in.
- ☐ b. Ephemeral personalization is based on demographics or similar characteristics rather than on actions or purchases.
- ☐ c. Ephemeral personalization is based on the products preferred by people like you -- your neighborhood of users.
- ☒ d. Ephemeral personalization is based on your current navigation or market basket, but not a long-term profile of your preferences.

1  
point

3. In the Hacker News scoring algorithm, the net upvotes is raised to a small ( $\leq 1$ ) power. Why?

- ☐ a. To demote old items.
- ☒ b. To decrease the weight of later upvotes, so votes 900-1000 have less individual impact than 90-100.
- ☐ c. To keep overall-popular items from being pinned at the top of the page.
- ☐ d. Because the relative importance of votes and age differences shift as items get older

1  
point

4. Amazon.com has many recommender systems. Which of the following techniques did we NOT see in our tour of Amazon.com?

- ☒ a. Recommenders based on demographics such as age and zip code
- ☐ b. Product association recommenders based on the page currently viewed
- ☐ c. Recommenders based on your recent shopping history
- ☐ d. Recommenders based on a long-term profile of purchases and ratings

1  
point

5. Why might we prefer product-association recommenders to average-rating recommenders?

- ☐ a. Product associations are based on user ratings.
- ☐ b. Product associations involve fancy math.
- ☒ c. Product associations allow recommendations that are relevant to a current context.
- ☐ d. Product associations are personalized to a user's full history, while average-ratings are not.

1  
point

6. When is “inverse document frequency” least useful as part of a content-filtering recommender?
- ☒ a. When attributes that apply only rarely are not all that helpful in making decisions (e.g., extras in a movie).
  - ☐ b. When certain items have many more attributes than other items.
  - ☐ c. When certain items are much more popular than other items.
  - ☐ d. When users are most interested in receiving recommendations for less popular items.

1  
point

7. The vector space model is quite useful for modeling document needs (i.e., what a user requests in a query) or item preferences (i.e., attribute preferences), but it has some limitations. Which of the following is a serious limitation of the model?
- ☐ a. It only works in domains where liking is a yes-no decisions; it can't handle degrees of preference.
  - ☒ b. It limits preferences to a linear combination of attributes -- it can't specify that you either want Tom Hanks and Meg Ryan together, or neither of them, but not one without the other.
  - ☐ c. It results in profiles that are nearly impossible to explain to an ordinary user because they are based on complex combinations of attributes that don't make intuitive sense.
  - ☐ d. It cannot produce top-n lists -- only predictions for individual item preferences.

1  
point

8. Which of these statements DOES NOT describe the Entrée Style Recommenders?
- ☒ a. They build a model of user preferences that can be used to provide personalized recommendations.
  - ☐ b. They require a substantial collection of information about the items being recommended.

- ☐ c. They provide an interface that allows the user to refine recommendations by requesting items that differ in a certain way from the current recommendation.
- ☐ d. They don't use individual users' ratings of the items anywhere in the recommendation process.

1  
point

9. The Herlocker explanations paper explored a variety of explanation interfaces, but it did have one key mistake. What was that mistake?

- ☐ a. It forgot to use some of the better explanations available.
- ☐ b. The authors didn't realize that some of the explanations were really just made up data.
- ☒ c. The paper really didn't measure usefulness of explanations; it measured persuasiveness of those explanations instead.
- ☐ d. The authors didn't recruit enough test users to get any statistically significant results.

1  
point

10. Which of the following would most indicate a situation where user-user collaborative filtering would be strongly preferable to content-based filtering?

- ☐ a. There are lots of items to recommend, and relatively few users.
- ☐ b. Only implicit ratings are available; users won't provide explicit ratings.
- ☐ c. Most users have rated a core set of popular items, though they have different tastes on that core set.
- ☒ d. The items being recommended don't have good attributes or keywords to describe them (e.g., user-submitted children's drawings without tags).

1  
point

11. Which of the following types of users have been the source of data for making recommendations in recommender systems?

- ☐ a. People with similar tastes to the target user.
- ☐ b. All system users who have expressed opinions.
- ☐ c. Experts whose opinions were solicited for the site.
- ☒ d. All of the above, sometimes in different systems, sometimes in the same system

1  
point

12. Resnick talked about resistance of collaborative filtering recommender systems to attacks from fake accounts (called sybils). Which of these statements about this problem is true?

- ☒ a. In order to be resistant to attacks from more sybils, you lose predictive power from genuine raters.
- ☐ b. The only way to be resistant to attacks from sybils is to trick them into rating fake movies that reveal that they aren't real users.
- ☐ c. If a collaborative filtering recommender is robust against attack from  $n$  sybils, a similar content-filtering recommender will only be resistant to attack from  $(n/2 - 1)$  sybils.
- ☐ d. There is unfortunately no way to bound the theoretical damage associated with a specific number of sybils -- no matter what you do, three clever sybils can inflict unlimited damage.

Submit Quiz

