twu38-HW6

Tianqi Wu

12/5/2019

Problem 1

Web search ads are identified by matching the keywords in the user search queries. Contextual ads are identified by matching the theme of contents that users read when browsing. The process of finding relevant ads are similar since ads are both labeled with keywords and identified based on users' interest. For search ads, users actively engage in the search process and the ads are usually more effective in this case. Whereas, for contextual ads, users engage in the process of reading contents during browsing and the ads are less effective. Contextual ads are less focused and less intentional. Also, much more text in the contents are needed to match the contextual ads.

Problem 2

Click through rate (CTR) measures the chance that users click the ads when they see the ads. For web search ads, users search with keywords and ads are displayed for the keywords. Hence, CTR for web search ads are indeed meaningful and relatively high. However, for contextual ads, it is harder to precisely determine users' intention and theme of the contents. Therefore, CTR for contextual ads are less meaningful and usually very low since it is difficult to match the ads with users' interest.

Problem 3

1.

Impedance mismatch refers to the problem that there is low intersection between vocabularies of the webpages and ads, even when they are related. It exists since ads tend to be more topic restricted than webpages and advertisers tend to use keywords describing broader concepts.

2.

To solve impedance mismatch problem, we can either generalize the terms in webpages or find more detailed terms for ads. To generalize the terms in webpages, we could expand the triggering pages with new terms by finding similar documents. To get more detailed terms for ads, we could inspect the landing pages linked by the ads.

3.

The keywords in triggering pages are too detailed and expanding them by introducing more generalized terms would help. The keywords in ads are too generalized and using the contents linked by the ads would help. These two additional sources of evidence are distinct and complementary. Hence, they provide valuable information which could give the highest precision.

Problem 4

Part 1

2

Interests:

- Food and drink : 75%
- News and Entertainment : 35%
- Business and Industry: 30%

3

Categories: 25%

4

The accuracy of my Facebook Ad profile is very low. The inference of categories is especially bad and I am not interested in most of entries. The reason might be that I am not a frequent user and I do not have many activities on Facebook. My interests change through time and the entries may be outdated. However, I like food very much and I want to try most of the food. Hence, there are a lot of correct entries for food and drink. These entries are likely to trigger by my activities on Facebook and beacons of Facebook.

Part 2

2

After entering different features, I narrowed down the potential reach to be less than 1000 people. The helpfulness of features from high to low: Language > Age > Location > Gender. Language is very helpful since my targeting friend is a Chinese. Chinese with location in US narrowed down the potential reach a lot. Also, since I know the exact age, It also helps a lot. Gender is relatively weak in this case since it could only half the potential reach.

3

My friend would probably not see my ad even the audience size is very low (<1000). Because the accuracy of my Facebook Ad profile is very low (25%) and the entries in categories are different from the entries in interests. Most of the entries are not related to my features (Chinese). Likewise, some features that I know about my friend may not be known by Facebook and the audience may exclude my friend. Hence, I believe that the probability that my friend would see the ad is very low.