Personalized click prediction in sponsored search

Cheng, H., & Cantú-Paz, E. (2010)

Presenter: Sarah Kim

2018.04.25

1. Introduction

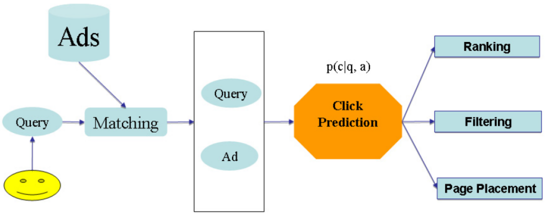

Figure 1 : Overview of sponsored search system

- Sponsored search: an Internet advertising system that generates most of the revenue of search engines by presenting targeted ads along with the search results.

1. Introduction

- Conventional approach
  - Use a machine learned model based on user-independent features to predict the click probability of ads.
  - Features for a machine learned model:
    1. the similarity of the query to the text of the ads;
    2. the historical performance of ads;
    3. contextual information (ex. time of day or day of the week).
  - ✓ This model will predict the same probability of click for every user.
- The objective of this paper is to design of pernalized click prediction models by developing new user-related features.

2. Click prediction

- Settings
  - $\mathcal{D} = \{(f(q_j, a_j), c_j)\}_{j=1}^{n}$: n training samples
  - $f(q_j, a_j) \in \mathbb{R}^d$: the d-dimensional feature space for query-ad pair j, $c_j \in \{-1, +1\}$: corresponding class label ($+1$: click or -1:non-click)
- Given a query q and ad a, the maximum entropy model (ME) formulates the click probability as follows:

$$p(c|q,a) = \frac{1}{1 + \exp(\sum_{i=1}^{d} w_i f_i(q,a))},$$

where $\mathbf{w} \in \mathbb{R}^d$ are weight parameters. Given the training set $\mathcal{D}$, $\mathbf{w}$ estimated as

$$\hat{\mathbf{w}} = \operatorname{argmax} \left[\sum_{j=1}^{n} \log(p(c_j|q_j, a_j)) + \log(p(\mathbf{w})) \right].$$

2. Click prediction

2.1 Features

- Features derived from the historical performace of ads:
  - Click-through rate (CTR) at position r :

    $$CTR_r(q, a) = \frac{c_r(q, a)}{i_r(q, a)},$$

    where $i_r(q, a)$ is the number of times that query q and ad a were shown together at position r, and $c_r(q, a)$ is the number of times those impression were clicked.
  - Clicks over expected clicks (COEC):

    $$COEC(q, a) = \frac{\sum_{r=1}^{R} c_r(q, a)}{\sum_{r=1}^{R} i_r(q, a) \times CTR_r},$$

    where CTR_r is the average CTR for each position r computed over all queries and ads and the denominator can be viewed as the expected clicks (ECs).
  - NV: total number of ad views
  - NCLI: total number of ad clicks

- Feature quantization:
  - In this work, the features are transformed into the log form and then quantized using simple K-means clustering.
  - We introduce binary indicator features for each cluster, and use these binary features as inputs to the ME model.
  - For new ads or new queris, we also use a binary indicator feature to indicate that a certain value is missing.
- Feature conjuctions:
  - To model relationships among features, create feature conjunctions by taking the cross product of the binary indicators for pairs of features.
  - We select the features to be conjoined using domain knowledge.

3. User click analysis

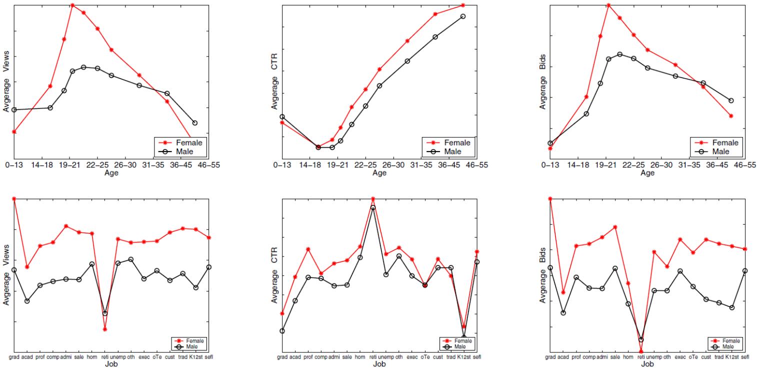

Figure 2 : User click, view and bid distributions with regard to demographic background

4. Personalized click prediction

- Let $\mathcal{D} = \{(f(q_j, a_j, u_j), c_j)\}_{j=1}^{n}$ represent the new training set, where each sample j represents a click or non-click event when ad a_j is presented to user u_j for query q_j.
- We develop a new click prediction function as:

$$p(c|q, a, u) = \frac{1}{1 + \exp(\sum_{i=1}^{d} w_i f_i(q, a, u))},$$

where $\mathbf{w} \in \mathbb{R}^d$ are weight parameters. Given the training set $\mathcal{D}$, $\mathbf{w}$ estimated as

$$\hat{\mathbf{w}} = \operatorname{argmax} \left[\sum_{j=1}^{n} \log(p(c_j|q_j, a_j, u_j)) + \log(p(\mathbf{w})) \right].$$

5. User features

1. Demographic group features
   - We partition users into demographic segments based on age, gender, marriage status, interests, job status, and occupation.
2. User-specific features
   - User-specific features capture individual user's interaction with the ads shown in the search result page.
   - $UCOEC_u = \frac{CTR \text{ of user } u}{\text{Average } CTR \text{ of } u\text{'s group}}$,
     where the average CTR is calculated for each user group and users are grouped together based on the total number of ads they have seen.
   - We can also derive other user-specific click feedback features (EC, COEC, NV, NCLI) at the user, user-query, and user-ad levels.

6. Experiments

- The training and testing data were sampled from the Yahoo! sponsored search traffic logs for a period of 2 months.
- Online performance comparison: for online testing we selected the model with user and user-ad features.

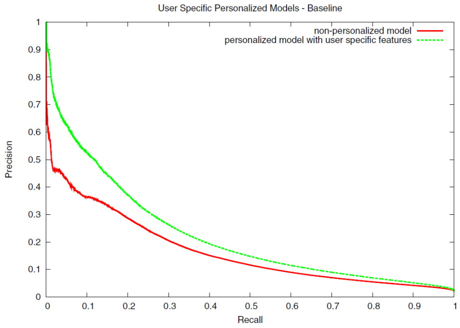

Figure 3 : Online performance of the personalized model.