Supplemental Material for Unbiased Learning to Rank

Qingyao Ai  
CICS, UMass Amherst  
Amherst, MA, USA  
aiqy@cs.umass.edu

Yiqun Liu  
DCST, Tsinghua University  
Beijing, China  
yiqunliu@tsinghua.edu.cn

Jiaxin Mao  
DCST, Tsinghua University  
Beijing, China  
maojiaxin@gmail.com

W. Bruce Croft  
CICS, UMass Amherst  
Amherst, MA, USA  
croft@cs.umass.edu

ABSTRACT
This is the supplemental material for the tutorial of Unbiased Learning to Rank: Theory and Practice. It includes some basic equations used in the tutorial. Please refer to the tutorial slides for detailed information.

ACM Reference format:  
https://doi.org/10.1145/3234944.3234980

1 MOTIVATION
The goal of unbiased learning to rank is to learn an unbiased ranker with noisy or biased feedback signals. In this case, we consider the real relevance between a document and the information need of a user (which is usually expressed with a query) as the true ranking signals we want to capture, and clicks on each query-document pair as the observed feedback signals we have.

The notation used in this tutorial is listed in Table 1. Formally, given the true relevance \( r_i \) for document \( x_i \), an optimal unbiased learning-to-rank model \( S(x_i, \theta) \) can be computed as

\[
\theta^* = \arg \min_{\theta} L(S) = \arg \min_{\theta} \int_{Q} \sum_{x_i \in \pi_q, r_i = 1} \Delta(S(x_i, \theta), r_i | \pi_q) \, dQ
\]

If we directly replace \( r_i \) with the observed click \( c_i \) on the documents, then the empirical model we get in practice is

\[
\hat{\theta} = \arg \min_{\theta} \hat{L}(S) = \arg \min_{\theta} \int_{Q} \sum_{x_i \in \pi_q, c_i \in \{1, 0\}} \Delta(S(x_i, \theta), c_i | \pi_q) \, dQ
\]

Therefore, our goal is to find a method that let the empirical model \( \hat{\theta} \) learned from clicks converge to the optimal model \( \theta^* \) learned from true relevance signals.

---

Table 1: A summary of notations used in this paper.

| \( Q, Q, q \) | The universal set of queries \( Q \), a sample set \( Q \) and a query instance \( q \). |
| \( S, \theta, E, \phi \) | A ranking system \( S \) parameterized by \( \theta \) and a propensity model \( E \) parameterized by \( \phi \). |
| \( \pi_q, x, i, y \) | A ranked list \( \pi_q \) produced by \( S \) for \( q \), a document \( x \) on the \( i \)th position in \( \pi_q \) and its label \( y \). |
| \( o_1, r_1, c_i \) | The sets of Bernoulli variables that represent whether a document \( x_i \) is observed (\( o_1 \)), perceived as relevant (\( r_1 \)) and clicked (\( c_i \)) in a search session. |
| \( O^s, P^s, R^s \) | The vectors of Bernoulli variables \( o_i, r_i, c_i \) on document \( x_i \) on a specific document \( x_i \) in a search session \( s \). |
| \( \Delta(x_i, y_i | \pi_q) \) | The loss of system \( S \) on a specific document \( x_i \) based on its label \( y_i \). |
| \( L(S), I(S, q) \) | The overall loss of system \( S \) (\( L(S) \)) and the local loss of system \( S \) on \( q \) (\( I(S, q) = \sum_{s \in \pi_q} y_i \Delta(x_i, y_i | \pi_q) \)). |

2 CLICK MODELS

2.1 Examination Hypothesis
The Examination Hypothesis assumes that a user will click a document if and only if she has examined it (\( o_1 = 1 \) and got attracted by the document (\( r_1 = 1 \)):

\[
c_i = 1 \iff o_i = 1 \text{ and } r_i = 1
\]

Based on this hypothesis, we have:

\[
P(c_i = 1) = P(o_i = 1) \times P(r_i = 1)
\]

We usually assume that \( P(r_i = 1) \) only depends on the document \( x \) and query \( q \):

\[
P(r_i = 1) = \alpha x q
\]

Different click models make different assumptions on users’ examination behavior, and therefore, have different implementation of \( P(o_i = 1) \).

2.2 MLE for Cascade Model
The log-likelihood of click models with parameters \( \Theta \) on the train sessions \(\mathcal{S}\) is

\[
LL(S, \Theta) = \sum_{s \in \mathcal{S}} \log P(c_1^s, c_2^s, \ldots, c_M^s | \Theta)
\]
$M_s$ is the number of documents retrieved by query $q$, usually for Web search, $M_s = 10$. The maximum-likelihood estimation of $\Theta$ is

$$\hat{\Theta} = \arg \max_{\Theta} LL(S, \Theta) \quad (7)$$

For cascade model [1], because $P_{CM}(o_k = 1|c_{k-1} = 1) = 0$ we know that for a search session $s$ that has a click on rank $l_s$:

$$o_{l_s+1} = o_{l_s+2} = \cdots = o_M = 0$$

Because $P_{CM}(o_k = 1|o_{k-1} = 0) = 0$ and $c_l = 1$ implies $o_l = 1$ and $r_l = 1$, we have:

$$o_1 = o_2 = \cdots = o_1 = 1$$

Therefore:

$$LL_{CM}(S, \Theta) = \sum_{s \in S} \log \Pi_{k=1}^{M_s} P(c_k^{(s)} = 1|x_q^{(s)} = 1 - P(c_k^{(s)} = 1)|^{1-c_k^{(s)}}$$

$$= \sum_{s \in S} \log \Pi_{k=1}^{M_s} (1 - \alpha_{s_k} q_s) \cdot \alpha_{s_k} q_s,$$

$$= \sum_{s \in S} \sum_{k=1}^{l_s} |i(k < l_s) log(1 - \alpha_{s_k} q_s) + I(k = l_s) log(\alpha_{s_k} q_s)|$$

$$\quad (8)$$

Here $S_{xq}$ is the set of search sessions initiated by query $q$ and containing document $x$.

With respect to each $\alpha_{s_q}$:

$$LL_{CM}(\alpha_{s_q}) = \sum_{s \in S_{xq}} \sum_{k=1}^{l_s} |i(k < l_s, x_k = x, q_s = q) log(1 - \alpha_{s_q})$$

$$+ \sum_{s \in S_{xq}} \sum_{k=1}^{l_s} |I(k = l_s, x_k = x, q_s = q) log(\alpha_{s_q})$$

$$\quad \text{+ Constant}$$

$$\quad (9)$$

Let $\frac{\partial LL_{CM}(\alpha_{s_q})}{\partial \alpha_{s_q}} = 0$, we can derive the MLE for $\alpha_{s_q}$:

$$\hat{\alpha}_{s_q} = \frac{ESS_{s,s_q}(1)}{ESS_{s,s_q}(1) + ESS_{s,s_q}(0)} \quad (10)$$

Where

$$ESS_{s,s_q}(1) = \sum_{s \in S_{xq}} \sum_{k=1}^{l_s} |i(k = l_s, x_k = x, q_s = q)$$

$$ESS_{s,s_q}(0) = \sum_{s \in S_{xq}} \sum_{k=1}^{l_s} |i(k < l_s, x_k = x, q_s = q)$$

$$\quad (11)$$

2.3 EM for UBM

The log-likelihood function of UBM [2] is

$$LL_{UBM}(S, \Theta) = \sum_{s \in S} \log P(C^{(s)}|\Theta)$$

$$= \sum_{s \in S} \sum_{O^{(s)}, R^{(s)}} \log P(C^{(s)}, O^{(s)}, R^{(s)}|\Theta)$$

$$\quad (12)$$

Due to the latent variables $O^{(s)}$ and $R^{(s)}$, the log-likelihood function is difficult to compute and optimize directly. Therefore, we alter to use the Expectation-Maximization (EM) algorithm to solve the parameter estimation problem of UBM. The EM algorithm iteratively optimize the Q-function:

$$Q_{UBM}(\Theta|\Theta') = \sum_{s \in S} E_{O^{(s)}, R^{(s)}|C^{(s)}, \Theta'} log P(C^{(s)}, O^{(s)}, R^{(s)}|\Theta)$$

$$= \sum_{s \in S} E_{O^{(s)}, R^{(s)}|C^{(s)}} \log \left( \frac{P(C_k^{(s)}|o_k^{(s)}, r_k^{(s)}) P(o_k^{(s)}|C_k^{(s)})}{P(r_k^{(s)}|\sigma_{xq}, q_k)} \right)$$

$$\quad (13)$$

In E-step, we will compute $P(O^{(s)}, R^{(s)}|C^{(s)}, \Theta')$ with old parameters $\Theta$. In M-step, we maximize $Q_{UBM}(\Theta|\Theta')$ w.r.t. $\Theta$ to update the parameters. E-steps and M-steps are taken iteratively until convergence, i.e., $\Theta = \Theta'$.

For UBM, $\Theta = \langle \{y_{rd}\}, \{a_{s_{q}}\} \rangle$. For each $a_{s_{q}}$:

$$Q(a_{s_{q}}|\Theta') = \sum_{s \in S_{xq}} I(C_s^{(s)} = 1) P(x_s^{(s)} = 1|c_s^{(s)} = 1, \Theta') log(a_{s_{q}})$$

$$+ I(c_s^{(s)} = 0) P(x_s^{(s)} = 1|c_s^{(s)} = 0, \Theta') log(a_{s_{q}})$$

$$+ I(c_s^{(s)} = 0) P(x_s^{(s)} = 0|c_s^{(s)} = 0, \Theta') log(1 - a_{s_{q}})$$

$$\quad + \text{Constant}$$

$$\quad (14)$$

In E-step, we compute:

$$P(r_s^{(s)} = 1|c_s^{(s)} = 1, \Theta') = 1$$

$$P(r_s^{(s)} = 1|c_s^{(s)} = 0, \Theta') = \alpha_{s_{q}} y_{rd}$$

$$P(r_s^{(s)} = 0|c_s^{(s)} = 0, \Theta') = \frac{1 - \alpha_{s_{q}}}{1 - \alpha_{s_{q}} y_{rd}}$$

$$\quad (15)$$

In M-step, we maximize $Q(a_{s_{q}}|\Theta')$ w.r.t. $a_{s_{q}}$:

$$\frac{\partial Q(a_{s_{q}}|\Theta')}{\partial a_{s_{q}}} = 0$$

$$a_{s_{q}} = \frac{1}{|S_{xq}|} \sum_{s \in S_{xq}} I(c_s^{(s)} = 1 - c_s^{(s)} \alpha_{s_{q}} y_{rd})$$

$$\quad (16)$$

For $y_{rd}$:

$$Q(y_{rd}|\Theta') = \sum_{s \in S_{rd}} I(c_s^{(s)} = 1) P(x_s^{(s)} = 1|c_s^{(s)} = 1, \Theta') log(y_{rd})$$

$$+ I(c_s^{(s)} = 0) P(x_s^{(s)} = 1|c_s^{(s)} = 0, \Theta') log(y_{rd})$$

$$+ I(c_s^{(s)} = 0) P(x_s^{(s)} = 0|c_s^{(s)} = 0, \Theta') log(1 - y_{rd})$$

$$\quad + \text{Constant}$$

$$\quad (17)$$

In E-step, we compute:

$$P(o_s^{(s)} = 1|c_s^{(s)} = 1, \Theta') = 1$$

$$P(o_s^{(s)} = 1|c_s^{(s)} = 0, \Theta') = \frac{1 - \alpha_{s_{q}} y_{rd}^{'}}{1 - \alpha_{s_{q}} y_{rd}^{'}}$$

$$P(o_s^{(s)} = 0|c_s^{(s)} = 0, \Theta') = \frac{1 - y_{rd}'}{1 - \alpha_{s_{q}} y_{rd}^{'}}$$

$$\quad (18)$$
In M-step, we optimize $Q(y_{rd} | \Theta^t)$ w.r.t. $y_{rd}$:

$$\frac{\partial Q(y_{rd} | \Theta^t)}{\partial y_{rd}} = 0$$

$$y_{rd} = \frac{1}{|S_{rd}|} \sum_{s \in S_{rd}} \left[ x_s (1 - c_s (1 - a'_{xy}) y_{rd}' \right] (1 - a'_{xy} y_{rd}') \right] (19)$$

3 UNBIASED LEARNING TO RANK

3.1 Inverse Propensity Weighting

Suppose that the original click-based ranking loss of a ranking system $S$ for a query $q$ is

$$l(S, q) = \sum_{x_i \in \pi_q, r_i = 1} \Delta(x_i, c_i | \pi_q)$$

Then, the ranking loss with inverse propensity weighting is [4]

$$l_{IPW}(S, q) = \sum_{x_i \in \pi_q, c_i = 1} \Delta(x_i, c_i | \pi_q) / P(o_i = 1 | \pi_q)$$

(21)

3.1.1 Proof of Correctness. As discussed previously, the true ranking loss based on the true relevance between each query-document pair is computed as

$$l(S, q) = \sum_{x_i \in \pi_q, r_i = 1} \Delta(x_i, r_i | \pi_q)$$

(22)

Assuming that

$$c_i = 1 \Rightarrow o_i = 1, r_i = 1$$

then we can prove that optimizing the inverse propensity weighted ranking loss can converge to the true ranking loss as [3]:

$$\mathbb{E}_{o_i} [l_{IPW}(S, q)] = \sum_{x_i \in \pi_q, c_i = 1, r_i = 1} \frac{\Delta(x_i, r_i | \pi_q)}{P(o_i = 1 | \pi_q)}$$

$$= \sum_{x_i \in \pi_q, r_i = 1} \mathbb{E}_{o_i} [o_i] \cdot \frac{\Delta(x_i, r_i | \pi_q)}{P(o_i = 1 | \pi_q)}$$

(23)

3.1.2 Proof of Robustness. Let the probability that a user would click a relevant document after observing it be

$$P(c_i = 1 | r_i = 1, o_i = 1) = \epsilon_+$$

Let the probability that a user would click an irrelevant document after observing it be

$$P(c_i = 1 | r_i = 0, o_i = 1) = \epsilon_-$$

Also, assuming that

$$\epsilon_+ > \epsilon_-$$

and the loss on each document $x_i$ is a function of its ranking position $\text{rank}(x_i | S)$ and $r_i$ as

$$\Delta(x_i, r_i | S) = f(\text{rank}(x_i | S)) \cdot r_i$$

(24)

then we have [3]

$$\mathbb{E}[l_{IPW}(S_1, q)] > \mathbb{E}[l_{IPW}(S_2, q)]$$

$$\Rightarrow \mathbb{E}_{o_i} \left[ \sum_{x_i \in \pi_q, c_i = 1} \frac{f(\text{rank}(x_i | S_1)) - f(\text{rank}(x_i | S_2))}{P(o_i = 1 | \pi_q)} \right] > 0$$

$$\Rightarrow \mathbb{E}_{o_i} \left[ \sum_{x_i \in \pi_q} P(o_i = 1 | \pi_q) \delta f(x_i) \right] > 0$$

$$\Rightarrow \mathbb{E}_{o_i} \left[ \sum_{x_i \in \pi_q} (\epsilon_+ r_i + \epsilon_- (1 - r_i)) \delta f(x_i) \right] > 0$$

(25)

$$\Rightarrow \mathbb{E}_{o_i} \left[ \sum_{x_i \in \pi_q} (\epsilon_+ - \epsilon_-) r_i \cdot \delta f(x_i) \right] > 0$$

$$\Rightarrow \mathbb{E}_{o_i} \left[ \sum_{x_i \in \pi_q} \delta f(x_i) \cdot r_i \right] > 0$$

$$\Rightarrow \mathbb{E}[l(S_1, q)] > \mathbb{E}[l(S_2, q)]$$

3.2 Propensity Estimation

3.2.1 Correctness of Online Randomization. We can prove that, in online result randomization, the expected value of click on position $k \in \pi_q$ is proportional to the real examination propensity [5]:

$$\mathbb{E}[c_k] = \int_{(q, x, \pi_q), i = k} P(c_i = 1 | \pi_q) dP(q, x, \pi_q)$$

$$= \int_{(q, x, \pi_q), i = k} P(o_i = 1) \cdot P(r_i = 1 | \pi_q) dP(q, x, \pi_q)$$

$$= P(o_k = 1) \cdot \int_{(q, x, \pi_q), i = k} P(r_i = 1 | \pi_q) dP(q, x, \pi_q)$$

$$\Rightarrow P(o_k = 1)$$

(26)

REFERENCES


