Extra Experiments on Propensity-Weighted Ranking SVM using arXiv Logs

Yuning Mao
Shanghai Jiao Tong University, Shanghai, China
morningmoni@sjtu.edu.cn

ABSTRACT

Implicit feedback is of great value for optimizing human interactive system, and yet the feedback could be severely biased due to the ways it was collected. To cope with this problem, [3] proposed an Empirical Risk Minimization (ERM) approach to Learning-to-Rank that guarantees consistent learning even for biased click data. Their approach is based on a counterfactual model adapted from causal inference, and they claimed that it does not require that the same query occurs multiple times as commonly assumed by click models. They derived a Propensity-Weighted Ranking SVM for discriminative learning from implicit feedback. Based on their approach, we did extra experiments using the click data collected from arXiv and showed that the results were consistently improved when propensities were estimated properly but became worse when propensities were misspecified.

1. INTRODUCTION

Although annotated test collections are of great help for batch training of retrieval systems, the efforts of collecting them are huge. Collecting implicit feedback from user behavior is hence useful and relatively economical. Existing approaches for Learning-to-Rank (LTR) from implicit feedback, in particular clicks on search results, have several limitations. For examples, they're usually biased in ignoring the presented order [4]; They require the same query to appear multiple times; Randomization of rankings decreases ranking quality. Due to the drawbacks mentioned above, [3] presented a theoretically principled and empirically effective approach for learning from observational implicit feedback. By drawing on counterfactual estimation techniques from causal inference [2], They developed a provably unbiased estimator for evaluating ranking performance using biased feedback data, based on which they proposed a Propensity-Weighted Empirical Risk Minimization (ERM) approach to LTR named Propensity-Weighted Ranking SVM.

In this paper, we did extra experiments using the click data collected from arXiv and showed that the results were consistently improved when propensities were estimated properly but became worse when propensities were misspecified.

2. UNBIASED LEARNING TO RANK

One problem of utilizing implicit feedback is that it's distorted by presentation bias and it's not missing at random [5]. [3] adopted the following counterfactual model to tackle it. They assumed that relevance is binary, $r_i(y) \in \{0, 1\}$, and they are interested in optimizing the rank of the relevant documents.

$$U(x_i, y) = \sum_{y \in \mathbf{y}} rank(y|\mathbf{y}) \cdot r_i(y) \tag{1}$$

For each incoming query instance $x_i \sim P(x)$, there exists a true vector of relevances r_i. Let o_i denote the 0/1 vector indicating which relevance values were revealed. Furthermore, denote with $Q(o_i(y) = 1|x_i, \overline{y}_i)$ the marginal probability of observing the relevance $r_i(y)$ of document y for query x_i, if the user was presented the ranking $\overline{y}_i$.

Then they derived an unbiased estimate of $U(x_i, y)$ for any new ranking y via the inverse propensity scoring (IPS) estimator [2] [1] [6]:

$$\begin{aligned}\hat{U}_{IPS}(x_i, y) =& \sum_{y \in \mathbf{y}} \sum_{o_i y \in \{0,1\}} \frac{o_i(y)}{Q(o_i(y) = 1|x_i, \overline{y}_i)} rank(y|\mathbf{y}) \cdot r_i(y) \\ =& \sum_{y \in \mathbf{y}} \frac{1}{Q(o_i(y) = 1|x_i, \overline{y}_i)} rank(y|\mathbf{y}) \cdot r_i(y)\end{aligned}$$

Note that it's weighted only by the propensities of observed relevant documents. This is an unbiased estimate of $U(x_i, y)$ for any new y, assuming that $Q(o_i(y) = 1|x_i, \overline{y}_i) > 0$ for all relevant $\mathbf{y}$, since

$$\begin{aligned}E[\hat{U}_{IPS}(x_i, y)] =& E_{o_i}[\sum_{y \in \mathbf{y}} \sum_{o_i(y) \in \{0,1\}} \frac{o_i(y)}{Q(o_i(y) = 1|x_i, \overline{y}_i)} \cdot \\ & rank(y|\mathbf{y}) \cdot r_i(y)] \\ =& \sum_{y \in \mathbf{y}} \frac{1 \cdot Q(o_i(y) = 1|x_i, \overline{y}_i)}{Q(o_i(y) = 1|x_i, \overline{y}_i)} rank(y|\mathbf{y}) \cdot r_i(y) \\ =& \sum_{y \in \mathbf{y}} rank(y|\mathbf{y}) \cdot r_i(y) \\ =& U(x_i, y)\end{aligned}$$

Since this estimate is unbiased for each x_i and y, the following empirical risk of system S

$$\hat{U}_{IPS}(S) = \sum_{x_i} \sum_{y \in S(x_i)} \frac{1}{Q(o_i(y) = 1 | x_i, \overline{y}_i)} rank(y|S(x_i)) \cdot r_i(y) \tag{2}$$

is also unbiased.

$$E[\hat{U}_{IPS}(S)] = U(S) \tag{3}$$

This means that we can perform empirical risk minimization using this propensity-scored empirical risk

$$\hat{S} = argmin_{S \in \mathbf{S}} \{\hat{U}_{IPS}(S)\} \tag{4}$$

given a sample of queries x_i, the partially-revealed relevances r_i as indicated by o_i, and the propensities $Q(o_i(y) = 1 | x_i, \overline{y}_i)$ under the rankings $\overline{y}_i$ presented during logging.

3. PROPENSITY ESTIMATION

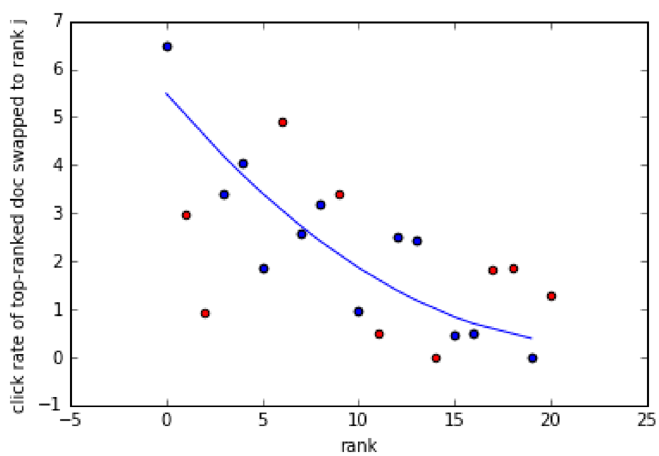

Figure 1: The fitting curve of click rate of top-ranked documents swapped to rank j. Outliers denoted by red nodes are removed

We collect new click data where we intervene in the following way. Before presenting the ranking, we take the top-ranked document and swap it with the document at rank $j \in [1, 21]$ with same probability. We then keep track of how often the two swapped documents in their new positions were clicked after each move.

We assume a position discount model with $P(click|pos) = P(click|rel, obs)P(obs|pos)P(rel|pos)$, where $P(click|rel, obs)$ is some constant. We can then estimate $P(obs|pos)$, namely, the propensity from the decay of clicks of the formerly top-ranked document as it is moved to a lower position. Since $P(rel|pos)$ is the same at all ranks for the swapped top-ranked document, namely $P(rel|pos) = P(rel|pos = 1)$, and $P(click|rel, obs)$ is a constant as well, $Q = P(obs|pos) \propto P(click|pos)$. We can then use the new click data to estimate propensity.

We first estimate the propensity by the method introduced above. We remove some outliers and then fit a quadratic function as shown in 1. We assign propensity of rank greater than 21 equal to that at rank 21.

In addition, we also tried other rough propensity estimation methods:

1. $Propensity = 1/j$, where j is the rank of document.
2. $Propensity = 1/(\#click+1)$. Roughly estimate propensity at a rank as the number of clicks on it.
3. Slightly Different Propensity: $Propensity = 1$ for rank 1-10, $Propensity = 1/1.1$ for rank 11-20, $Propensity = 1/1.2$ for rank > 20.
4. Rough Propensity Estimation: $Propensity = 1$ for rank 1-10, $Propensity = 1/2$ for rank 11-20, $Propensity = 1/3$ for rank > 20.
5. Counterfactual Propensity Estimation: $Propensity = 1/3$ for rank 1-10, $Propensity = 1/2$ for rank 11-20, $Propensity = 1$ for rank > 20.

4. PROPENSITY-WEIGHTED RANKING SVM

We briefly restate the Propensity-Weighted Ranking SVM [3] here. The Propensity-Weighted Ranking SVM implements the propensity-scored ERM for Learning to Rank. It is based on the Ranking SVM. For each query-document pair (x_i, y_i) that is clicked, the propensity q_i of it is recorded. Multiple clicks in the same query are regarded as separate training examples. The optimization objective is as follows:

$$\hat{w} = argmin_w \frac{1}{2} w \cdot w + \frac{C}{n} \sum_1^n \frac{1}{q} \sum_y \xi_{ij} \tag{5}$$

This optimizes an upper bound on the IPS estimated empirical risk, since each line of constraints corresponds to the rank of a relevant document (minus 1). We can solve this type of QP very efficiently via a one-slack formulation.

5. REAL-WORLD EXPERIMENT

5.1 Data Description

To evaluate the model empirically, we use click data on arXiv. The features are well-defined in this experiment. The total number of features is 1000. The data we collected contains 3205 valid clicks (the clicks on documents whose features were stored) and 4664 different queries. The numbers of clicks on different positions are shown in 2.

5.2 Parameter Selection

For the selection of C, We start from a small C and times 2 each time to find the C that has best performance in terms of Average Rank of Positive Examples on validation set. Then We use it on the test set.

5.3 Same Propensity Estimation method in training set and test set

For training, we learn a ranking function using the propensity-weighted ranking SVM with the estimated propensities. We also learn another ranking function using the naive ranking SVM. We use 60% as training data, 30% as test data, and 10% as validation data. We first use same propensity estimation methods in both training set of the weighted ranking SVM and test set, and assign unweighted propensity in the training set of normal ranking SVM. The average ranking of positive examples on test set are shown in 4.

As we can see, the weighted ranking SVM is always slightly better (but not significantly) than the normal ranking SVM. But this may result from the fact that the propensities in the

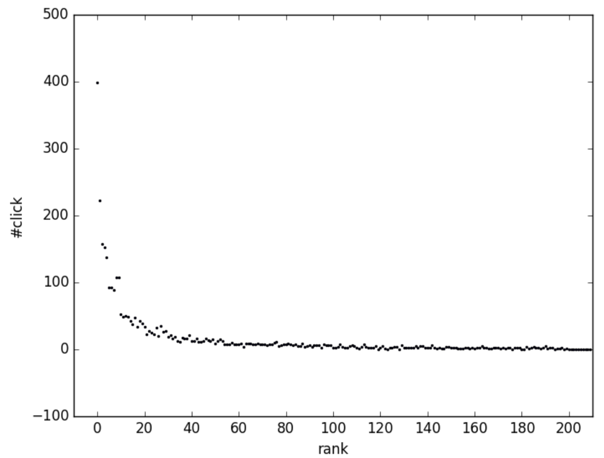

Figure 2: Click count of different positions

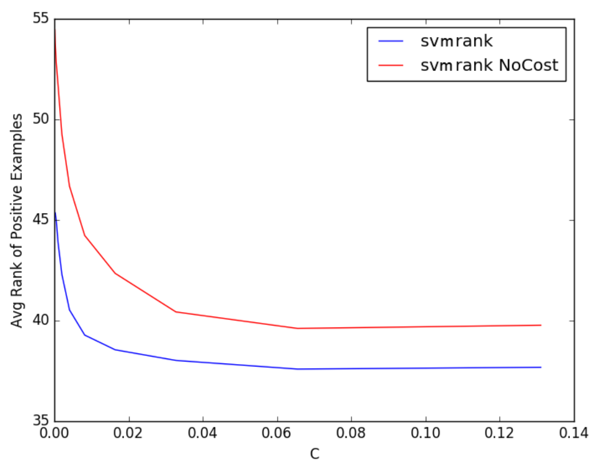

Figure 3: The learning curve of weighted Ranking SVM and normal Ranking SVM on validation set

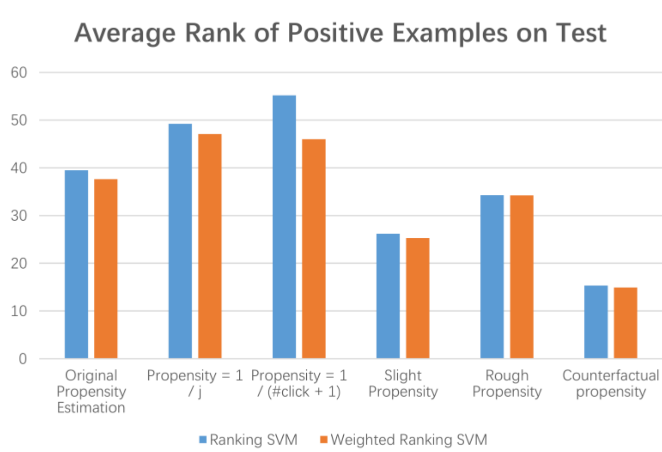

Figure 4: Average Rank of Positive Examples on Test when using Same Propensity Estimation methods

training set (for weighted ranking SVM) and in the test set are estimated using same method. Since weighted ranking SVM utilizes more information about the data, itâĂŹs supposed to outperform the normal ranking SVM. Therefore, we then tried using different propensity estimation methods in training set and test set, namely, rough propensity estimation in the training set and more accurate estimation in the test set.

5.4 Rough Propensity Estimation in training set with More Accurate Estimation in test set

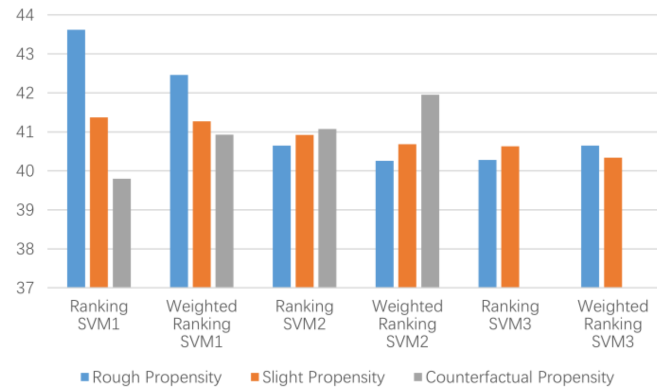

Figure 5: Average Rank of Positive Examples on Test using Different Propensity Estimation methods

Next we tried using different propensity estimation methods in training set and test/validation set, in which we suppose the original propensity estimation method was true (used in the test set).

The results also look good when the propensities in the training set decrease as ranks increase (as in the test set), as can be seen when Rough Propensity and Slight Propensity are used. This indicates that even if the estimated propensities are not precise, considering the propensity can still improve performance, if the propensity estimation method is relatively reasonable.

Nevertheless, the performance became worse the other way around, when Counterfactual Propensity is used, which means that the model degrades gracefully under misspecified propensities. The average ranking of positive examples on test set are shown in 5.

6. CONCLUSIONS

In this paper, we review the method in [3] that utilizes biased feedback from users and how it makes the feedback unbiased for the use of learning-to-rank. More specifically, they employed the theory in causal inference to consider the propensities of observing relevances of documents in rankings. They proved that this approach is unbiased when propensities are taken into account. They proposed the Propensity-Weighted Ranking SVM. Based on their approach, we used click data on arXiv to verify the model and ideas in [3]. By empirical evaluation, we show that the Propensity-Weighted Ranking SVM outperforms the normal Ranking SVM, even when the estimated propensities are not accurate; But misspecified propensities may lead to worse performance.

7. ACKNOWLEDGMENTS

I want to thank Prof. Thorsten Joachims for his instructions and time. I also thank Adith for helping me understand the paper.

8. REFERENCES

[1] D. G. Horvitz and D. J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685, 1952.

[2] G. W. Imbens and D. B. Rubin. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press, 2015.

[3] T. Joachims. Unbiased learning-to-rank with biased feedback. In *WSDM*, 2017.

[4] T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, and G. Gay. Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search. *ACM Transactions on Information Systems (TOIS)*, 25(2):7, 2007.

[5] R. J. Little and D. B. Rubin. *Statistical analysis with missing data*. John Wiley & Sons, 2014.

[6] P. R. Rosenbaum and D. B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.