Asymmetric Tri-Training for Debiasing Missing-Not-At-Random Explicit Feedback

Yuta Saito (https://usaito.github.io/)
Undergraduate Student
Tokyo Institute of Technology
Collaborative Filtering Approach

Learn users’ preferences on items from observed ratings

<table>
<thead>
<tr>
<th>Items</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>-</td>
</tr>
</tbody>
</table>

\[ \approx \]

Matrix Factorization

\[
\begin{align*}
U & \approx V
\end{align*}
\]
True and observed rating distributions are different..

**true rating distribution**  \[\xrightarrow{\text{experimentally estimated}}\]  **observed rating distribution**

**Selection bias**
- past recommendation policy
- users’ self-selection

(a) Yahoo! Survey Rating Distribution

(b) Yahoo! Base Rating Distribution

Figure 2. Marlin, B., Zemel, R. S., Roweis, S., and Slaney, M. Collaborative filtering and the missing at random assumption. In UAI, 2007.
In summary,

The selection bias issue breaks the assumption of machine learning

**Train and Test (true) distributions are different in recommender systems**

Addressing the selection bias is essential in the evaluation and learning of recommender systems offline

Let’s analyze the issues using statistical tools!
Performance measure in the “ideal world”

Given a set of predicted ratings for all user-item pairs \( \hat{R} = \{ \hat{R}_{u,i} \}_{(u,i)} \)

\[
\mathcal{L}(\hat{R}) = \frac{1}{U \cdot I} \sum_{u,i} \text{loss}(R_{u,i}, \hat{R}_{u,i})
\]

(local loss (squared loss, absolute loss))

empirical mean under uniform user-item distribution
Estimating the “ideal world” performance is critical

Recommender model’s parameters are updated based on estimated loss

\[ \hat{L}(\hat{R}) = ? \]

- observed world
- estimation

\[ L(\hat{R}) = \frac{1}{U \cdot I} \sum_{u,i} \text{loss}(R_{u,i}, \hat{R}_{u,i}) \]

- ideal world
Modeling Missing Mechanisms

We use the following observation indicator to model missing mechanisms:

$$O_{u,i} = \begin{cases} 
1 & (R_{u,i} \text{ is observed}) \\
0 & \text{(otherwise)} 
\end{cases}$$

Thus training data is:

$$\mathcal{O} = \{ (u, i, R_{u,i}) : O_{u,i} = 1 \}$$
“Naive” Estimator for the “Ideal World”

The naive estimator is

the empirical mean of local loss over the observed data

$$\mathcal{L}_{naive}(\hat{R}) = \frac{1}{|O|} \sum_{(u,i): O_{ui}=1} \text{loss}(R_{u,i}, \hat{R}_{u,i})$$

observed data

most recommender systems attempt to optimize this naive loss
Naive estimator is “biased”

The expectation of the naive estimator fails to approximate the ideal world

\[
\mathbb{E}_O \left[ \widehat{\mathcal{L}}_{\text{naive}}(\widehat{R}) \right] = \mathbb{E}_O \left[ \frac{1}{|O|} \sum_{u,i} O_{u,i} \cdot \text{loss} \left( R_{u,i}, \widehat{R}_{u,i} \right) \right] \\
= \frac{1}{|O|} \sum_{u,i} \mathbb{E}_{O_{u,i}} \left[ O_{u,i} \right] \cdot \text{loss} \left( R_{u,i}, \widehat{R}_{u,i} \right) \\
\neq \frac{1}{U \cdot I} \sum_{u,i} \text{loss} \left( R_{u,i}, \widehat{R}_{u,i} \right)
\]

\[\mathcal{L}(\widehat{R})\] biased!!
Inverse Propensity Score (IPS) Estimator for the “Ideal World”

IPS estimator removes the bias by **weighting local loss** by the inverse of the propensity score

\[
\hat{L}_{IPS}(\hat{R}) = \frac{1}{U \cdot I} \sum_{(u,i):O_{ui}=1} \frac{\text{loss}(R_{ui}, \hat{R}_{ui})}{P_{u,i}}
\]

*observed data*  
*propensity score*  

\[
P_{u,i} = \mathbb{E}[O_{u,i}]
\]
IPS estimator is “unbiased”

IPS estimator can approximate the ideal world in expectation

\[
\mathbb{E}_O \left[ \hat{\mathcal{L}}_{IPS}(\hat{R}) \right] = \mathbb{E}_O \left[ \frac{1}{U \cdot I} \sum_{u,i} O_{u,i} \cdot \frac{loss(R_{u,i}, \hat{R}_{u,i})}{P_{u,i}} \right]
\]

\[
= \frac{1}{U \cdot I} \sum_{u,i} \frac{\mathbb{E}O_{u,i} \left[ O_{u,i} \right]}{P_{u,i}} \cdot loss(R_{u,i}, \hat{R}_{u,i})
\]

\[
= \frac{1}{U \cdot I} \sum_{u,i} loss(R_{u,i}, \hat{R}_{u,i}) = \mathcal{L}(\hat{R})
\]

Should we really use IPS?
Issues with the IPS estimator

- **Bias issue**
  - To ensure IPS’s unbiasedness, the true propensity score is needed. But, it is hard to estimate the propensity due to users’ self-selection (uncontrollable by analysts)

- **Variance issue**
  - IPS estimator can have a huge variance when the observed data is highly sparse
Our proposal: Asymmetric Tri-Training

To overcome the issues with IPS, we propose a model-agnostic meta-learning algorithm called “asymmetric-tri training”

Asymmetric-tri training uses three base recommenders and consists of the following three steps

1. Pre-Training Step
2. Pseudo-Labeling Step
3. Final Prediction Step
Step 1: Pre-Training Step

Asymmetric-tri training has three base recommenders
At the pre-training step, we pre-train three base recommender

$A_1, A_2, A_3$

We can use any recommendation model at the pre-training step such as Naive MF, MF-IPS, Factorization Machines.
Step 2: Pseudo-Labeling Step

At this step, we create **reliable pseudo-ratings** by using A1 & A2

\[
\tilde{D} = \left\{ \left( u, i, \hat{R}_{u,i}^{(1)} \right) : \left| \hat{R}_{u,i}^{(1)} - \hat{R}_{u,i}^{(2)} \right| \leq \epsilon \right\}
\]

pseudo-rating dataset

threshold hyperparameter (should be tuned)
Step 3: Final Prediction Step

Further update the other predictor A3 by using pseudo ratings

\[ \ell_{pseudo} (\hat{R}^{(3)}, \hat{R}^{(1)}) = \frac{1}{|\tilde{D}|} \sum_{(u, i) \in \tilde{D}} \ell (\hat{R}_{u,i}^{(3)}, \hat{R}_{u,i}^{(1)}) \]

Outputs by A3 are used as the final predictions
Asymmetric-tri training consists of the following three steps:

1. **Pre-Training Step**
   - pretrain three base recommenders A1, A2, and A3

2. **Pseudo-Labeling Step**
   - obtain reliable pseudo ratings by using A1 and A2

3. **Final Prediction Step**
   - further update A3 by using pseudo rating dataset
Propensity-independent upper bound of the “ideal world” loss

\[ \mathcal{L}_{\text{ideal}}(\hat{R}, \hat{\mathcal{R}}) \leq \mathcal{L}_{\text{pseudo}} \left( \hat{R}, \hat{R}^{(1)} \right) + \mathcal{L}_{\text{ideal}}^{\ell} \left( \hat{R}^{(1)}, \hat{R}^{(2)} \right) + \ldots \]

(a) is minimized at the final prediction step
(b) is kept small (not minimized) at the pseudo-labeling step
Theoretical Interpretation (Section 4.2)

Propensity-independent upper bound of the “ideal world” loss

\[ \mathcal{L}_{\text{ideal}}(\hat{R}, \hat{R}) \leq \mathcal{L}_{\text{pseudo}} \left( \hat{R}, \hat{R}^{(1)} \right) + \mathcal{L}_{\text{ideal}}^{\ell} \left( \hat{R}^{(1)}, \hat{R}^{(2)} \right) + \ldots \]

Even if IPS-based models are used as A1 and A2, issues with IPS are expected to be removed.
Experiment: Datasets

We used the following **Yahoo! R3 and Coat** datasets especially suitable for the MNAR recommendation

(a) Yahoo! R3 ($KL-div = 0.470$)

(b) Coat ($KL-div = 0.049$)
Experiment: Datasets

Both datasets have different train-test distributions
(Bias of Yahoo! R3 is much more larger than that of Coat)

(a) Yahoo! R3 ($KL$-div = 0.470)

(b) Coat ($KL$-div = 0.049)
Experiment: Compared Method

We tested

Matrix factorization with Inverse Propensity Score using six different propensity score estimators w/ or w/o the “asymmetric tri-training (AT)”

= 12 methods
Experiment: Compared Method

Six propensity score estimators

= Five practical estimators and One ideal (NB; true) estimators

uniform propensity: \( \hat{P}_{*,*} = \frac{\sum_{u,i \in D} O_{u,i}}{|D|} \)

user propensity: \( \hat{P}_{u,*} = \frac{\sum_{i \in I} O_{u,i}}{\max_{u \in U} \sum_{i \in I} O_{u,i}} \)

item propensity: \( \hat{P}_{*,i} = \frac{\sum_{u \in U} O_{u,i}}{\max_{i \in I} \sum_{u \in U} O_{u,i}} \)

user-item propensity: \( \hat{P}_{u,i} = \hat{P}_{u,*} \cdot \hat{P}_{*,i} \)

NB (uniform): \( \hat{P}_{u,i} = \mathbb{P}(R = R_{u,i} | O = 1) \mathbb{P}(O = 1) \)

NB (true): \( \hat{P}_{u,i} = \frac{\mathbb{P}(R = R_{u,i} | O = 1) \mathbb{P}(O = 1)}{\mathbb{P}(R = R_{u,i})} \)

use only biased train data

uses some amount of test data
(proposed in the original paper)
## Experiment: Issues with IPS

### Observation 1:

**MF with IPS fails when uniform log data is unavailable**

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Propensity</th>
<th>MAE without AT</th>
<th>MAE with AT</th>
<th>MSE without AT</th>
<th>MSE with AT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yahoo! R3</td>
<td>uniform</td>
<td>1.133</td>
<td>0.981</td>
<td>1.907</td>
<td>1.452</td>
</tr>
<tr>
<td></td>
<td>user</td>
<td>1.062</td>
<td>0.945</td>
<td>1.712</td>
<td>1.350</td>
</tr>
<tr>
<td></td>
<td>item</td>
<td>1.142</td>
<td>0.978</td>
<td>1.940</td>
<td>1.458</td>
</tr>
<tr>
<td></td>
<td>user-item</td>
<td>1.162</td>
<td>0.991</td>
<td>1.979</td>
<td>1.513</td>
</tr>
<tr>
<td></td>
<td>NB (uniform)</td>
<td>1.170</td>
<td>1.010</td>
<td>1.954</td>
<td>1.511</td>
</tr>
<tr>
<td></td>
<td>NB (true)</td>
<td>0.797</td>
<td>0.765</td>
<td>1.055</td>
<td>1.014</td>
</tr>
</tbody>
</table>

-impractical estimator-
Experiment: benefit of AT

Observation 2:

AT improves the original MF-IPS especially with only biased log

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Propensity</th>
<th>MAE without AT</th>
<th>MAE with AT</th>
<th>MSE without AT</th>
<th>MSE with AT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yahoo! R3</td>
<td>uniform</td>
<td>1.133</td>
<td>0.981</td>
<td>1.907</td>
<td>1.452</td>
</tr>
<tr>
<td></td>
<td>user</td>
<td>1.062</td>
<td>0.945</td>
<td>1.712</td>
<td>1.350</td>
</tr>
<tr>
<td></td>
<td>item</td>
<td>1.142</td>
<td>0.978</td>
<td>1.940</td>
<td>1.458</td>
</tr>
<tr>
<td></td>
<td>user-item</td>
<td>1.162</td>
<td>0.991</td>
<td>1.979</td>
<td>1.513</td>
</tr>
<tr>
<td></td>
<td>NB (uniform)</td>
<td>1.170</td>
<td>1.010</td>
<td>1.954</td>
<td>1.511</td>
</tr>
<tr>
<td></td>
<td>NB (true)</td>
<td>0.797</td>
<td>0.765</td>
<td>1.055</td>
<td>1.014</td>
</tr>
</tbody>
</table>
Experiment: upper bound minimization by AT

Observation 3:

AT successfully minimizes the theoretical upper bound

Figure 3: Upper bound minimization performance of asymmetric tri-training
Experiment: “ideal world” loss minimization by AT

Observation 4:
AT successfully optimizes the “ideal world” performance

Figure 4: Improved performance on the test sets by asymmetric tri-training
Conclusion

- We proposed the model-agnostic meta-learning method called “asymmetric tri-training” for debiasing biased explicit feedback.

- The proposed method minimizes the propensity independent upper bound of the “ideal world” loss.

- Empirical results verified the issues with the original IPS and our theoretical analysis.
Thank you for listening!

email: saito.y.bj at m.titech.ac.jp
preprint: https://usaito.github.io/publications/
github: https://github.com/usaito/asymmetric-tri-rec-real