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

Yuta Saito (https://usaito.github.io/)

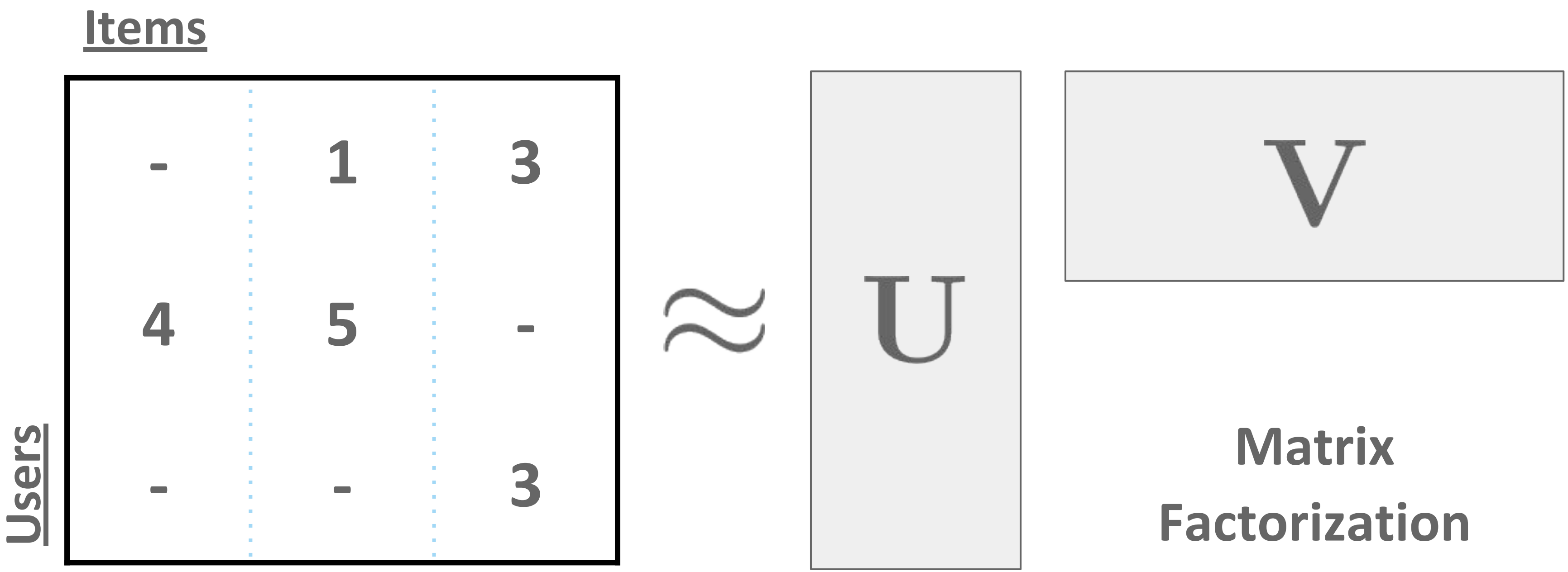
Undergraduate Student

Tokyo Institute of Technology

Tokyo Tech

Collaborative Filtering Approach

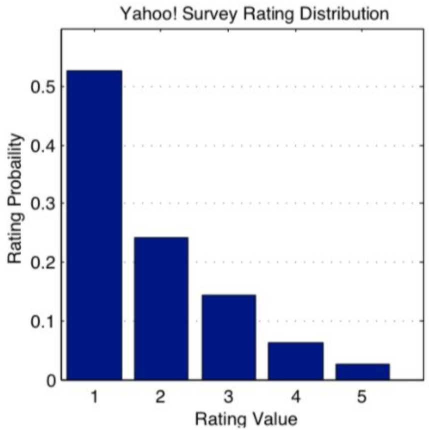
Learn users’ preferences on items from observed ratings



True and observed rating distributions are different..

true rating distribution ← experimentally estimated

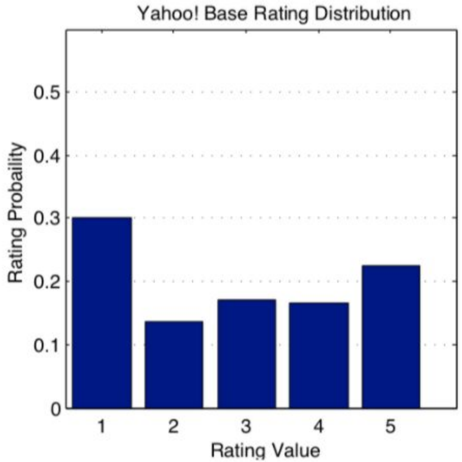
observed rating distribution



(a) Yahoo! Survey Rating Distribution

Selection bias

- past recommendation policy
- users' self-selection



(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.

(Marlin et al., UAI'07)

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 $\hat{R} = \left\{ \hat{R}_{u,i} \right\}_{(u,i)}$

$$\mathcal{L}(\hat{R}) = \underbrace{\frac{1}{U \cdot I} \sum_{u,i}}_{\text{empirical mean under uniform user-item distribution}} \underbrace{loss(R_{u,i}, \hat{R}_{u,i})}_{\text{local loss}}$$

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

observed world $\hat{\mathcal{L}}(\hat{R}) = $ **?**

↓ estimation ↓

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

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

$$\mathcal{O} = \{(u, i, R_{u,i}) : \underline{O_{u,i} = 1}\}$$

“Naive” Estimator for the “Ideal World”

The naive estimator is

the empirical mean of local loss over the observed data

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

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

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

$$= \frac{1}{|\mathcal{O}|}\sum_{u,i}\mathbb{E}_{O_{u,i}}\left[O_{u,i}\right]\cdot loss\left(R_{u,i},\hat{R}_{u,i}\right)$$

$$\neq \underbrace{\frac{1}{U\cdot I}\sum_{u,i} loss\left(R_{u,i},\hat{R}_{u,i}\right)}_{\mathcal{L}(\hat{R})}$$

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**

$$\widehat{\mathcal{L}}_{IPS}(\hat{R}) = \frac{1}{U \cdot I} \underbrace{\sum_{(u,i):O_{u\,i}=1}}_{\text{observed data}} \frac{loss(R_{u,i}, \hat{R}_{u,i})}{\underbrace{P_{u,i}}_{\text{propensity score}}}$$

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

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

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

Should we really use IPS?

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

unbiased!!

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 recommeders

$$A_1, A_2, A_3$$

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

Output by A1

Output by A2

$$\widetilde{\mathcal{D}} = \left\{ \left(u, i, \widehat{R}^{(1)}_{u,i}\right) : \left|\widehat{R}^{(1)}_{u,i} - \widehat{R}^{(2)}_{u,i}\right| \leq \epsilon \right\}$$

pseudo-rating dataset

threshold hyperparameter (should be tuned)

Step 3: Final Prediction Step

Further update the other predictor A3 by using pseudo ratings

Output by A3 → $\widehat{\boldsymbol{R}}^{(3)}$

Output by A1 (pseudo-ratings) → $\widehat{\boldsymbol{R}}^{(1)}$

$$\widehat{\mathcal{L}}^{\ell}_{pseudo}\left(\widehat{\boldsymbol{R}}^{(3)}, \widehat{\boldsymbol{R}}^{(1)}\right) = \frac{1}{|\widetilde{\mathcal{D}}|} \sum_{(u,i) \in \widetilde{\mathcal{D}}} \ell\left(\widehat{R}^{(3)}_{u,i}, \widehat{R}^{(1)}_{u,i}\right)$$

Outputs by A3 are used as the final predictions

Wrapping up: Asymmetric-tri Training

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

Theoretical Interpretation (Section 4.2)

Propensity-independent upper bound of the "ideal world" loss

$$\mathcal{L}_{ideal}(\mathbf{R}, \widehat{\mathbf{R}}) \leq \underbrace{\widehat{\mathcal{L}}_{pseudo}\left(\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{R}}^{(1)}\right)}_{(a)} + \underbrace{\mathcal{L}^{\ell}_{ideal}\left(\widehat{\boldsymbol{R}}^{(1)}, \widehat{\boldsymbol{R}}^{(2)}\right)}_{(b)} + \dots$$

(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

$$\mathcal{L}_{ideal}(\mathbf{R}, \widehat{\mathbf{R}}) \leq \underbrace{\widehat{\mathcal{L}}_{pseudo}\left(\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{R}}^{(1)}\right)}_{(a)} + \underbrace{\mathcal{L}^{\ell}_{ideal}\left(\widehat{\boldsymbol{R}}^{(1)}, \widehat{\boldsymbol{R}}^{(2)}\right)}_{(b)} + \dots$$

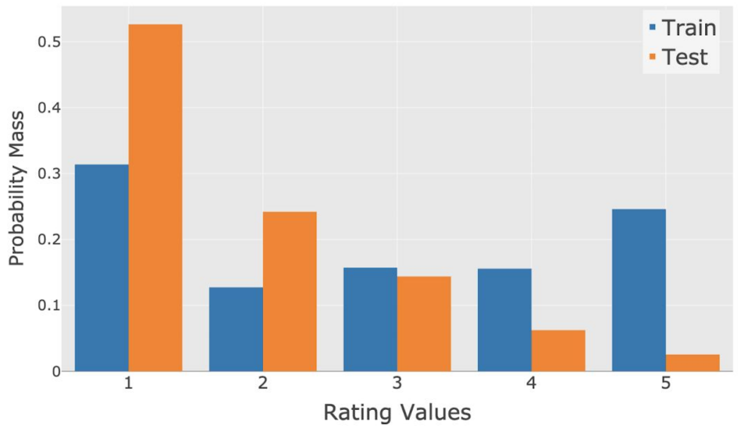
Even if IPS-based models are used as A1 and A2,

issues with IPS are expected to be removed

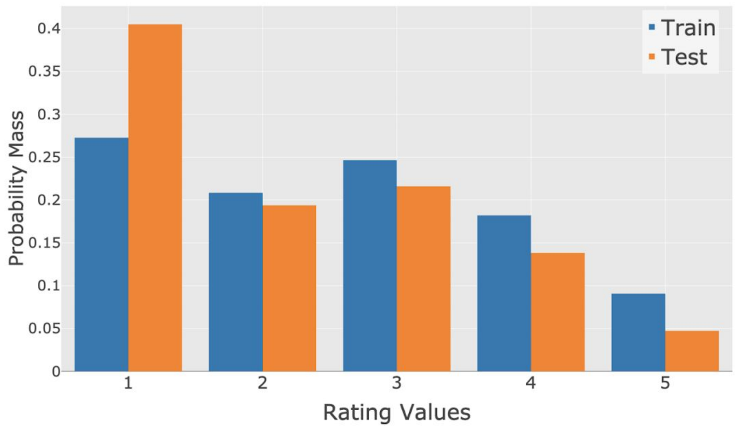
Experiment: Datasets

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)

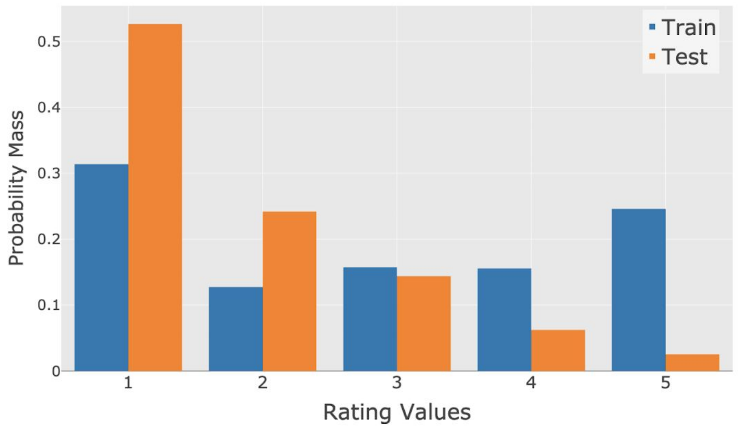


Experiment: Datasets

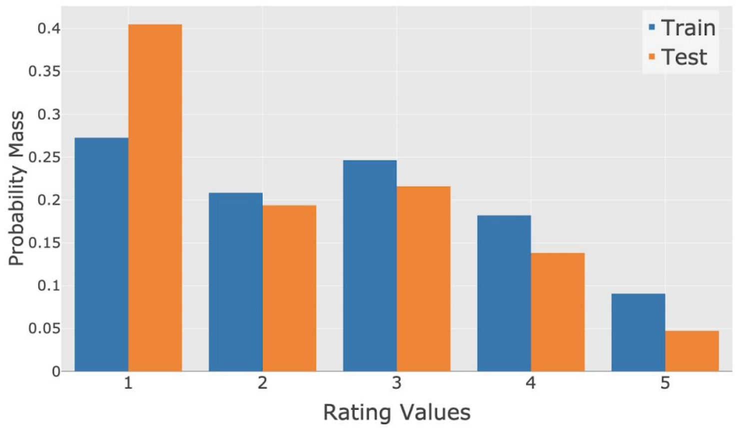
Both datasets have different train-test distributions

(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

$$\textit{uniform propensity} : \widehat{P}_{*,*} = \frac{\sum_{u,i \in \mathcal{D}} O_{u,i}}{|\mathcal{D}|}$$

$$\textit{user propensity} : \widehat{P}_{u,*} = \frac{\sum_{i \in \mathcal{I}} O_{u,i}}{\max_{u \in U} \sum_{i \in \mathcal{I}} O_{u,i}}$$

$$\textit{item propensity} : \widehat{P}_{*,i} = \frac{\sum_{u \in \mathcal{U}} O_{u,i}}{\max_{i \in I} \sum_{u \in \mathcal{U}} O_{u,i}}$$

$$\textit{user-item propensity} : \widehat{P}_{u,i} = \widehat{P}_{u,*} \cdot \widehat{P}_{*,i}$$

$$\textit{NB (uniform)} : \widehat{P}_{u,i} = \mathbb{P}(R = R_{u,i} \mid O = 1)\mathbb{P}(O = 1)$$

use only biased train data

$$\textit{NB (true)} : \widehat{P}_{u,i} = \frac{\mathbb{P}(R = R_{u,i} \mid O = 1)\mathbb{P}(O = 1)}{\mathbb{P}(R = R_{u,i})}$$

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

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

impractical estimator → NB (true)

Experiment: benefit of AT

Observation 2:

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

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

Experiment: upper bound minimization by AT

Observation 3:

AT successfully minimizes the theoretical upper bound

upper
bound of
the ideal loss

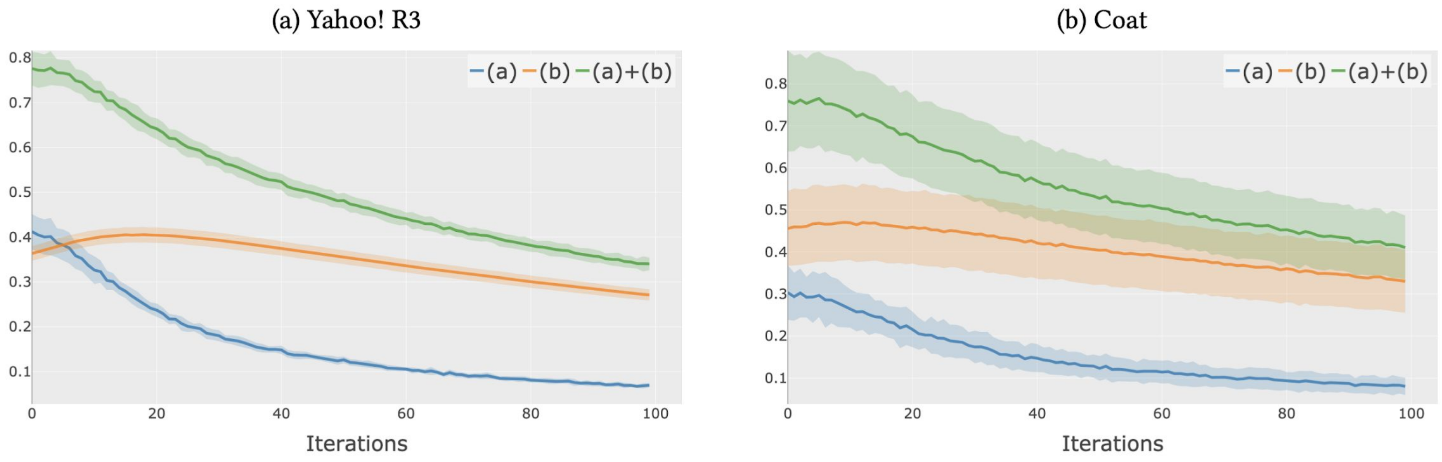


Figure 3: Upper bound minimization performance of *asymmetric tri-training*

Experiment: “ideal world” loss minimization by AT

Observation 4:

AT successfully optimizes the “ideal world” performance

the
ideal loss

$$\mathcal{L}_{ideal}(\widehat{\mathbf{R}})$$

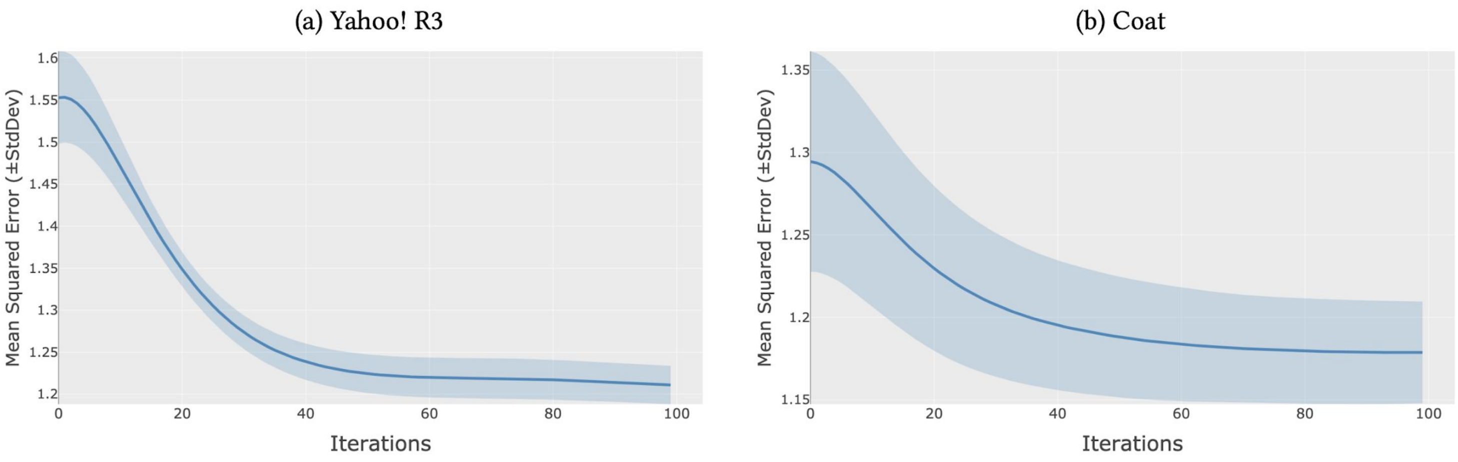


Figure 4: Improved performance on the test sets by *asymmetric tri-training*

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