Unsupervised Domain Adaptation Meets Offline Recommender Learning

Yuta Saito

Department of Industrial Engineering and Economics, Tokyo Institute of Technology

Tokyo Tech

Overview

- **Explicit (rating) feedback is often biased** by the past recommendation policies or user's self-selection in real-world recommender systems.
- **Propensity-based algorithms are the current promising solutions** to the challenge. However, **this approach needs some amount of test data to estimate the propensity score**.
- **We develop a new theoretical framework that is independent of the propensity score inspired by the framework of unsupervised domain adaptation**.
- Following the derived theoretical bound, we also propose a matrix factorization based algorithm called ***Domain Adversarial Matrix Factorization***. The proposed algorithm optimizes a mini-max objective via the adversarial learning.
- In the experiments, the proposed method outperformed the propensity-based models under situations where the true propensities are unknown.

Problem Setting and Existing Methods

- $\mathcal{U}$ is a set of users ($|\mathcal{U}| = m$), and $\mathcal{I}$ is a set of items ($|\mathcal{I}| = n$). $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ is a set of all user and items.
- $\boldsymbol{R} \in \mathbb{R}^{m \times n}$ be a true rating matrix, where each entry $R_{u,i}$ represents the true rating of u to i.
- Let $\boldsymbol{P} \in (0,1)^{m \times n}$ be a **propensity matrix**. Each entry of this matrix $P_{u,i}$ is the **propensity score of (u,i) representing the probability of the feedback being observed**.
- Let $\boldsymbol{O} \in \{0,1\}^{m \times n}$ be an **observation matrix** each entry $O_{u,i} \in \{0,1\}$ is a Bernoulli random variable with its expectation $\mathbb{E}[O_{u,i}] = P_{u,i}$.

The objective here is to develop an algorithm to obtain an optimal predicted rating matrix $\widehat{\boldsymbol{R}}$. To achieve this objective, **the following ideal loss function should be considered**:

$$\mathcal{L}^{\ell}_{\text{ideal}}\left(\widehat{\boldsymbol{R}}\right) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \ell\left(R_{u,i}, \widehat{R}_{u,i}\right) \tag{1}$$

Thus, **estimating the ideal loss from the observed rating feedback is critical to building a well-performing recommender offline**.

Given a feedback data $\boldsymbol{O}$, the most basic estimator is the **naive estimator**:

$$\widehat{\mathcal{L}}^{\ell}_{\text{naive}}\left(\widehat{\boldsymbol{R}}\,|\,\boldsymbol{O}\right) = \frac{1}{M} \sum_{(u,i)\in\mathcal{D}} O_{u,i} \cdot \ell\left(R_{u,i}, \widehat{R}_{u,i}\right) \tag{2}$$

Under the general MNAR settings, **this naive estimator actually has a bias**. Thus, it is undesirable to learn a recommendation algorithm.

On the other hand, one can derive an **unbiased estimator for the loss function of interest with the true propensity score (IPS estimator)** as follows:

$$\widehat{\mathcal{L}}^{\ell}_{\text{IPS}}\left(\widehat{\boldsymbol{R}}\,|\,\boldsymbol{O}\right) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} O_{u,i} \cdot \frac{\ell\left(R_{u,i}, \widehat{R}_{u,i}\right)}{P_{u,i}} \tag{3}$$

Theoretical Bound

The previous IPS estimator provides the unbiased loss function. However, for its unbiasedness, **this estimator needs true propensity matrix** that is generally unavailable in the real-world.

To alleviate the problem, we utilized the theory of **unsupervised domain adaptation** because of the following reasons.

- Similar to the propensity weighting, **unsupervised domain adaptation addresses the problem settings in which the feature distributions between the training and test sets are different**.
- Moreover, methods of unsupervised domain adaptation generally utilize distance metrics measuring dissimilarity between probability distributions and **does not depend on propensity weighting techniques**.

We derive **an upper bound of the ideal loss that is independent of the propensity score**.

Theorem 1. (Propensity-agnostic Upper Bound) Two observation matrices $\boldsymbol{O}_{\text{MCAR}} \sim \boldsymbol{P}_{\text{MCAR}}$ and $\boldsymbol{O}_{\text{MNAR}} \sim \boldsymbol{P}_{\text{MNAR}}$ having MCAR and MNAR missing mechanisms, respectively, and any finite hypothesis space of predictions $\mathcal{H} = \{\widehat{\boldsymbol{R}}_1, \ldots, \widehat{\boldsymbol{R}}_{|\mathcal{H}|}\}$ are given. The loss function ℓ is bounded above by a positive constant Δ. Then, for any hypothesis $\widehat{\boldsymbol{R}} \in \mathcal{H}$ and for any $\delta \in (0,1)$, the following holds with a probability of at least $1-\delta$

$$\mathcal{L}^{\ell}_{\text{ideal}}\left(\widehat{\boldsymbol{R}}\right) \leq \underbrace{\widehat{\mathcal{L}}^{\ell}_{\text{naive}}\left(\widehat{\boldsymbol{R}}\,|\,\boldsymbol{O}_{\text{MNAR}}\right)}_{\text{naive loss}} + \underbrace{d_{\mathcal{H}\Delta\mathcal{H}}\left(\boldsymbol{O}_{\text{MCAR}}, \boldsymbol{O}_{\text{MNAR}}\right)}_{\text{discrepancy measure}}$$
$$+ \frac{\Delta}{M}\left(\sqrt{\frac{|\mathcal{D}|}{2}\log\frac{4|\mathcal{H}|}{\delta}} + \sqrt{2|\mathcal{D}|\log\frac{8|\mathcal{H}|^2}{\delta}}\right) + \lambda \tag{4}$$

To construct the theoretical bound, we used **the discrepancy metric ($d_{\mathcal{H}\Delta\mathcal{H}}$) measuring the dissimilarity between the two propensity matrices**. Intuitively, it is the accuracy when we classify whether a give feedback data is from the MCAR or MNAR mechanism.

Proposed Objective and Algorithm

Inspired by the derived theoretical bound and algorithm in unsupervised domain adaptation, we propose the following **mini-max objective**.

$$\min_{\boldsymbol{U},\boldsymbol{V}} \max_{\theta_{\text{dom}}} \widehat{\mathcal{L}}_{\text{pred}}\left(\widehat{\boldsymbol{R}}\right) - \beta\widehat{\mathcal{L}}_{\text{dom}}\left(f;\theta_{\text{dom}}\right) \tag{5}$$

where $\widehat{\boldsymbol{R}} = \boldsymbol{U}\boldsymbol{V}^{\top}$, $\widehat{\mathcal{L}}_{\text{pred}}(\widehat{\boldsymbol{R}})$ is the **naive loss** defined in (1) and $\widehat{\mathcal{L}}_{\text{dom}}\left(f;\theta_{\text{dom}}\right)$ is called the **domain loss** and defined as follows.

$$\widehat{\mathcal{L}}_{\text{dom}}\left(f;\theta_{\text{dom}}\right) = \sum_{(u,i)\in\mathcal{D}_{\text{rare}}} \log\left(f\left(\boldsymbol{V}_i;\theta_{\text{dom}}\right)\right) + \sum_{(u,i)\in\mathcal{D}_{\text{pop}}} \log\left(1 - f\left(\boldsymbol{V}_i;\theta_{\text{dom}}\right)\right)$$

where $\mathcal{D}_{\text{rare}}$ is the set of all users and **rare** items, in contrast, $\mathcal{D}_{\text{pop}}$ is the set of all users and **popular** items. $\boldsymbol{V}$ is the item latent factor. The proposed ***Domain Adversarial Matrix Factorization*** algorithm optimizes its parameters iteratively using the Adam optimizer.

Experimental Results

We used the **MovieLens 1M (ML 1M), Yahoo! R3, and Coat datasets**. The Yahoo and Coat datasets originally have train and test data with the different missing mechanisms. For the ML 1M, we created three types of test sets with different missing mechanism from the training set. (Type 1 has a small, Type 2 has a medium, and Type 3 has a large difference from the training set.)

We compared the proposed algorithm with MF-IPS with several propensity estimators. **We used only the training sets to estimate the propensity score for MF-IPS because DAMF does not need any test data for training.** Note that MF-IPS (NB with true) used test datasets to estimate the propensity score. Thus, we report the results of this model just as a reference.

The results are reported in Table 1. There are two key observations.

- **MF-IPS, with propensity estimators using only training data, did not always outperform the vanilla MF (MF-IPS with uniform)**. The results suggest that MF-IPS is potentially an effective de-biasing method but is **highly sensitive to the way of propensity estimation**. In particular, using only MNAR training data for the propensity estimation does not lead to a well-performing recommender.
- **DAMF achieved significant performance gains on the three types of ML 1M and Yahoo! R3 dataset over the propensity-based MF models**. Moreover, the performance gain on the ML 1M type 3 is much larger than those on type 1 and 2. It outperformed 13.4 % in type 3, 11.9 % in type 2, and 9.5 % in type 1 over the best baselines.

	ML 1M (type 1)		ML 1M (type 2)		ML 1M (type 3)		Yahoo! R3		Coat	
	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE
MF (uniform)	0.905	1.282	1.024	1.585	1.201	2.012	1.150	2.044	0.803	1.091
MF-IPS (user)	0.936	1.353	1.114	1.814	1.330	2.339	1.128	1.973	0.815	1.127
MF-IPS (item)	0.938	1.368	1.075	1.739	1.292	2.277	1.101	1.973	**0.794**	**1.064**
MF-IPS (user-item)	0.957	1.410	1.097	1.789	1.297	2.289	1.110	2.009	0.935	1.444
MF-IPS (logistic)	0.908	1.293	1.031	1.605	1.187	1.979	1.157	2.075	0.856	1.249
DAMF (ours)	**0.890**	**1.215**	**0.968**	**1.416**	**1.110**	**1.744**	**0.991**	**1.566**	0.909	1.413
MF-IPS (NB with true)	0.845	1.143	0.741	0.908	0.562	0.592	0.796	1.095	0.793	1.107

Table 1: Experimental Results on the three datasets.

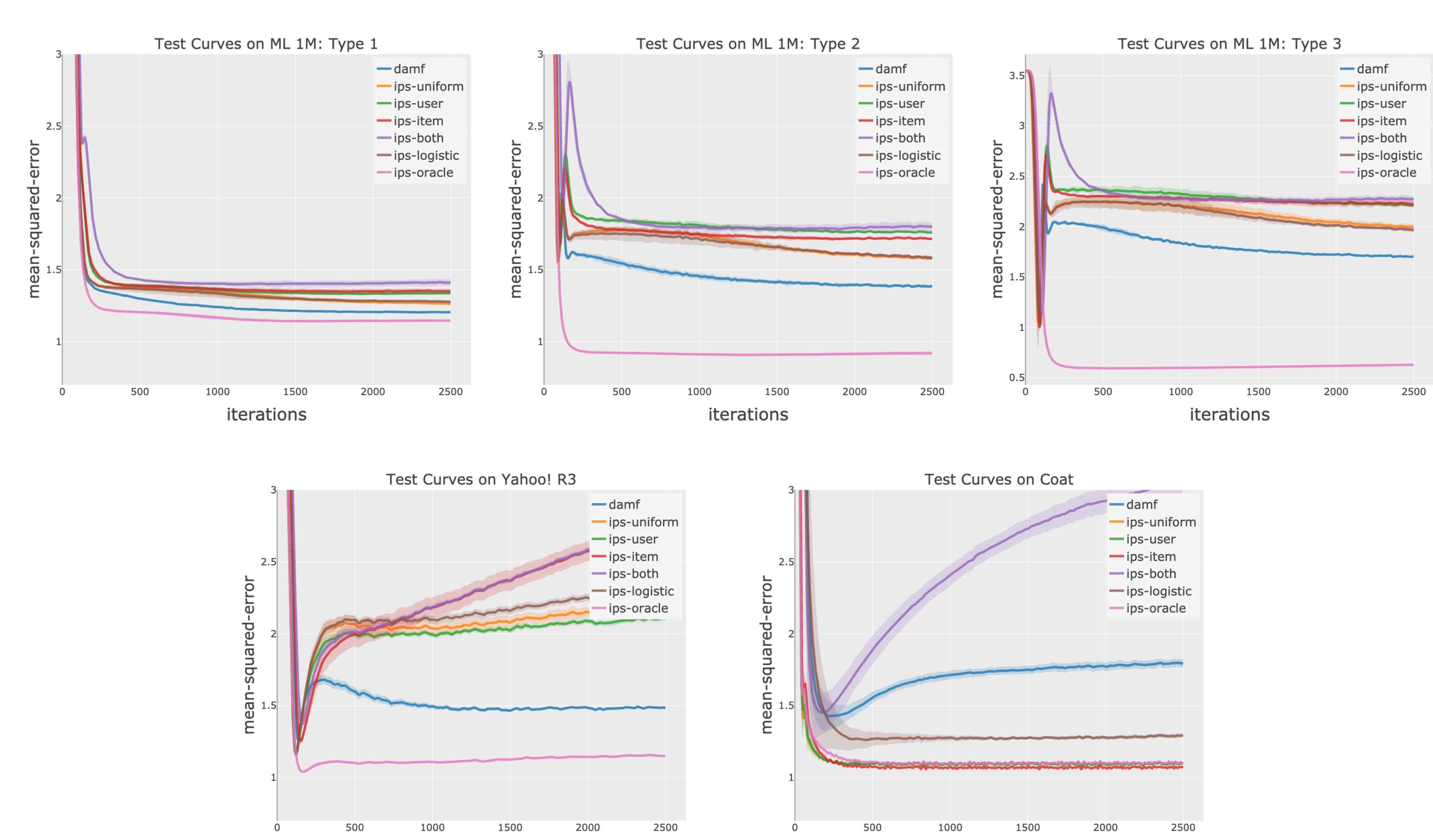

Fig. 1: Test MSEs vs. number of iterations on ML1M, Yahoo! R3 and Coat.

Future Work

The future research directions can be

- Bridging the gap between the theory and the algorithm.
- Applying the same theoretical framework to the implicit feedback setting.
- Empirical evaluations using online production environment.