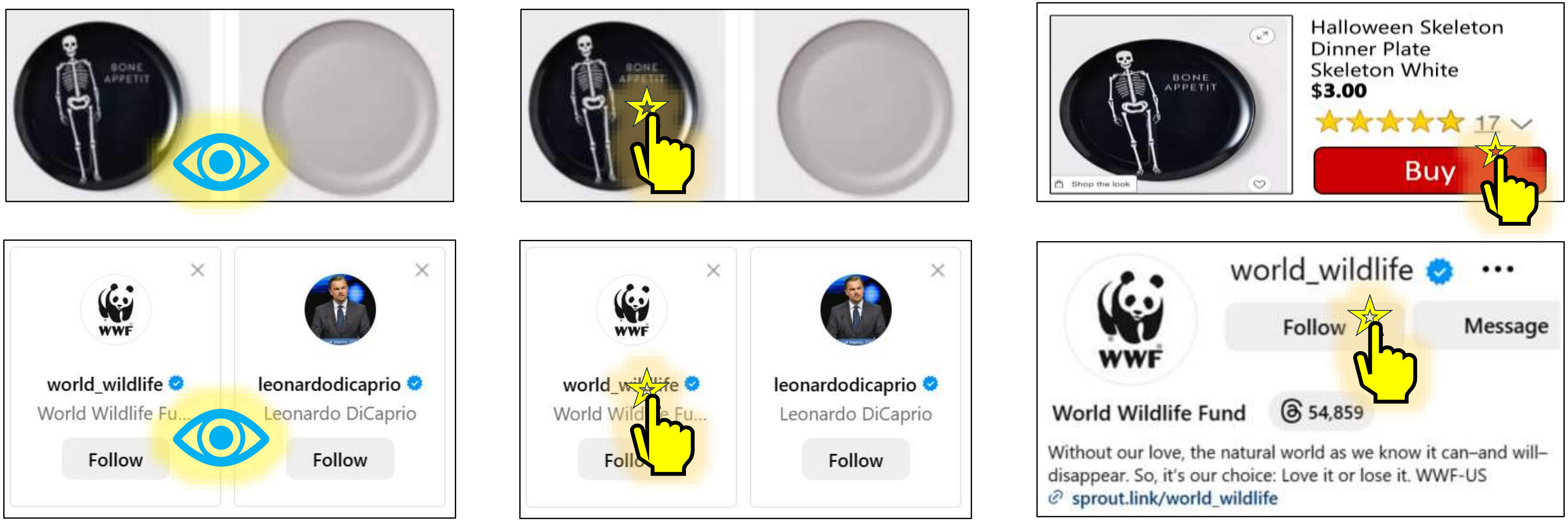
A Generalized Propensity Learning Framework for Unbiased Post-Click Conversion Rate Estimation

Yuqing Zhou, Tianshu Feng, Mingrui Liu, Ziwei Zhu

George Mason University

Recommender systems play an important role in people’s daily decision-making.

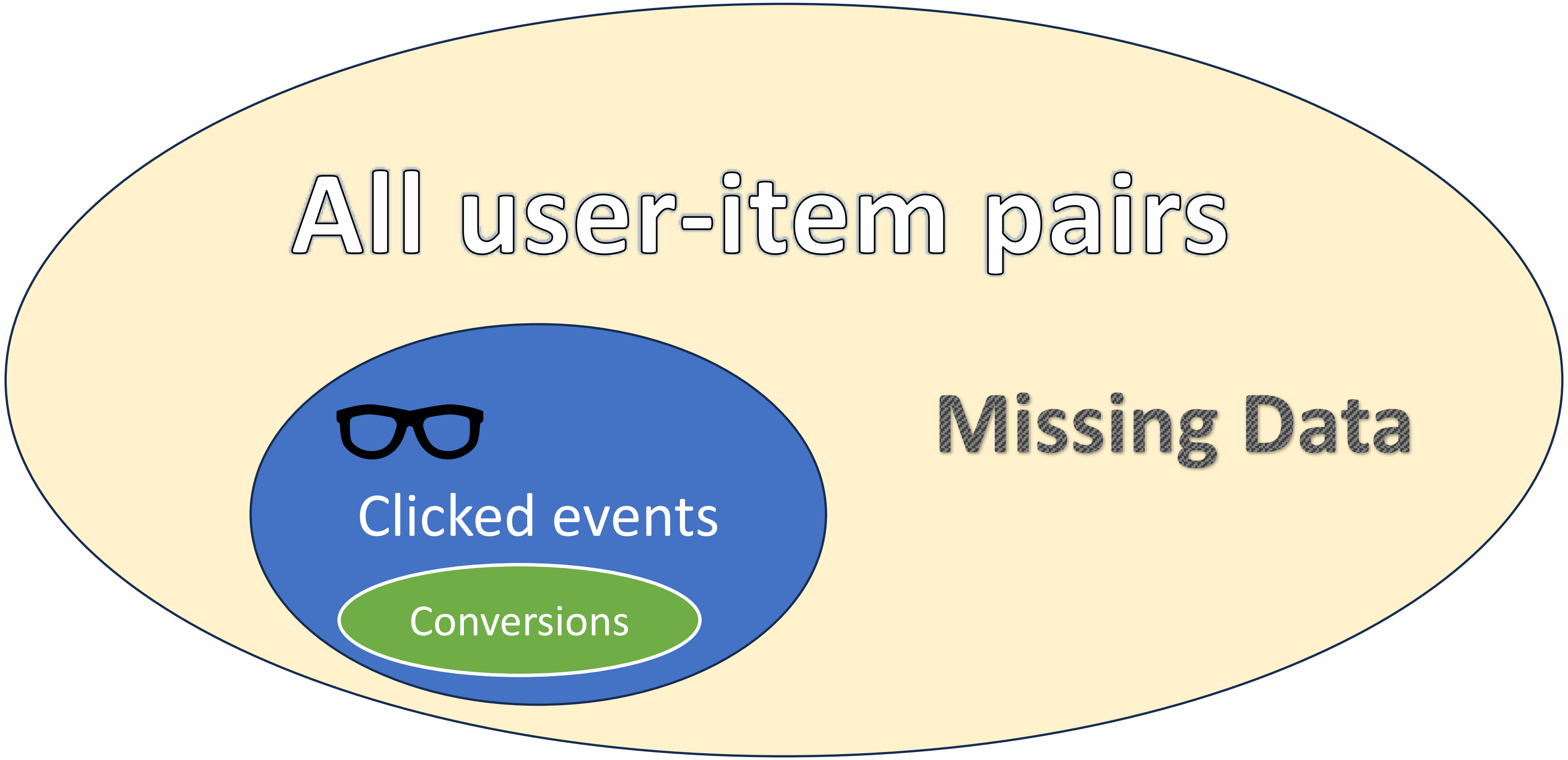
One important task for recommender systems：
Post-click Conversion Rate (**CVR**) prediction



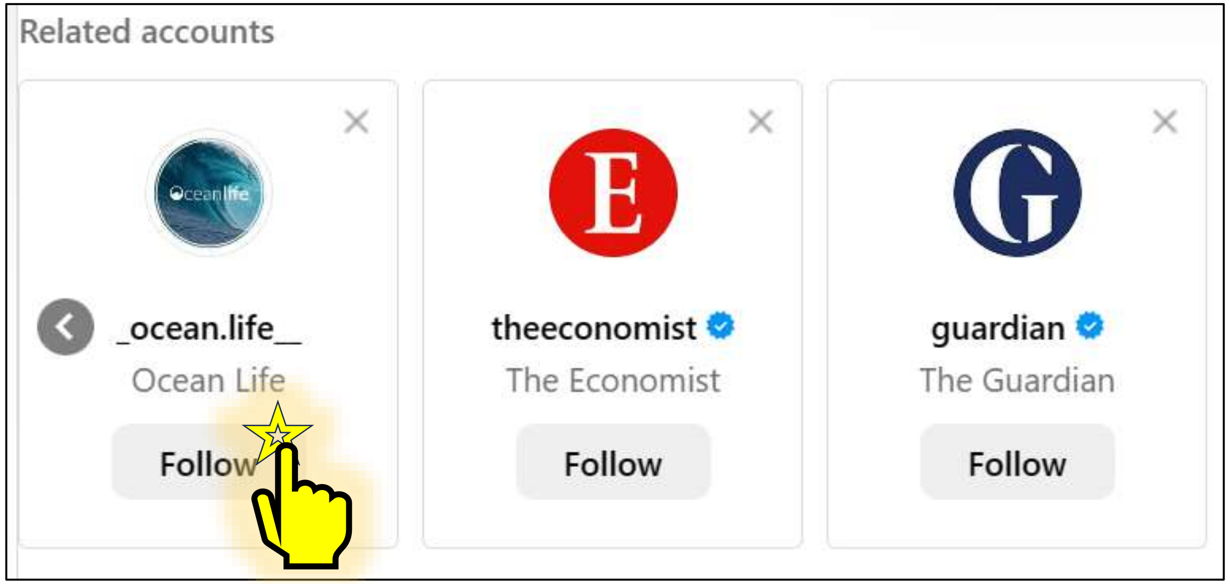
The **post-click CVR** indicates the **probability** of **obtaining a conversion** after a user clicks on an item, which is a **strong signal of user preference**.

The ideal post-click CVR estimation is conducted under the condition that all items are clicked by all users, and **all conversion results are observed**.

The fact is



The click data and the conversion data are missing-not-at-random (MNAR)



$P(click) = 0.7$ $P(click) = 0.2$ $P(click) = 0.1$

Assume the user will follow all three persons after clicking to see their homepages, i.e. $P(convert \mid click) = 1$.

In fact, users have different preferences toward different items when they decide what items to click. The probability a user clicks an item varies across users and items.

→ **Selection Bias**

Problem Formulation

- Ideally, given the fully observed conversion matrix, the loss function can be formulated as:

$$\mathcal{L}_{ideal}(\widehat{\boldsymbol{R}}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} e_{u,i}$$

where $e_{u,i}$ is the prediction error of $r_{u,i}$ and $\hat{r}_{u,i}$.

Problem Formulation

- Naive CVR estimation treats the missing data as missing-at-random.
- A naive estimator of the ideal loss function is

$$\mathcal{L}_{naive}(\widehat{\boldsymbol{R}}) = \frac{1}{|\mathcal{O}|} \sum_{(u,i)\in\mathcal{O}} e_{u,i} = \frac{1}{|\mathcal{O}|} \sum_{(u,i)\in\mathcal{D}} o_{u,i} e_{u,i}$$

$\mathcal{O}$ is the set of all clicked events.

Problem Formulation

- The probability of the item i being clicked by the user u is

$$p_{u,i} = P(o_{u,i} = 1) = E[o_{u,i}],$$

which is also called ***propensity*** or ***click-through-rate*** (CTR).

- $p_{u,i}$ varies across users and items, and the conversion results are missing-not-at-random(MNAR), so

$$E[\mathcal{L}_{naive}(\widehat{\boldsymbol{R}})] = \frac{1}{|\mathcal{O}|} \sum_{(u,i) \in \mathcal{D}} p_{u,i}\, e_{u,i} \neq \mathcal{L}_{ideal}(\widehat{\boldsymbol{R}})$$

Biased!

Existing methods to lower bias and variance of CVR estimation

- Inverse Propensity Score (IPS) Estimator

$$\mathcal{L}_{IPS}(\widehat{\boldsymbol{R}}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} e_{u,i}}{\hat{p}_{u,i}}$$

$\hat{p}_{u,i}$ is the estimation of $p_{u,i}$

- Doubly Robust (DR) Estimator and Its Variants (MRDR, DRMSE)

$$\mathcal{L}_{DR}(\widehat{\boldsymbol{R}}) = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} \hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i} - \hat{e}_{u,i})}{\hat{p}_{u,i}}$$

Motivation

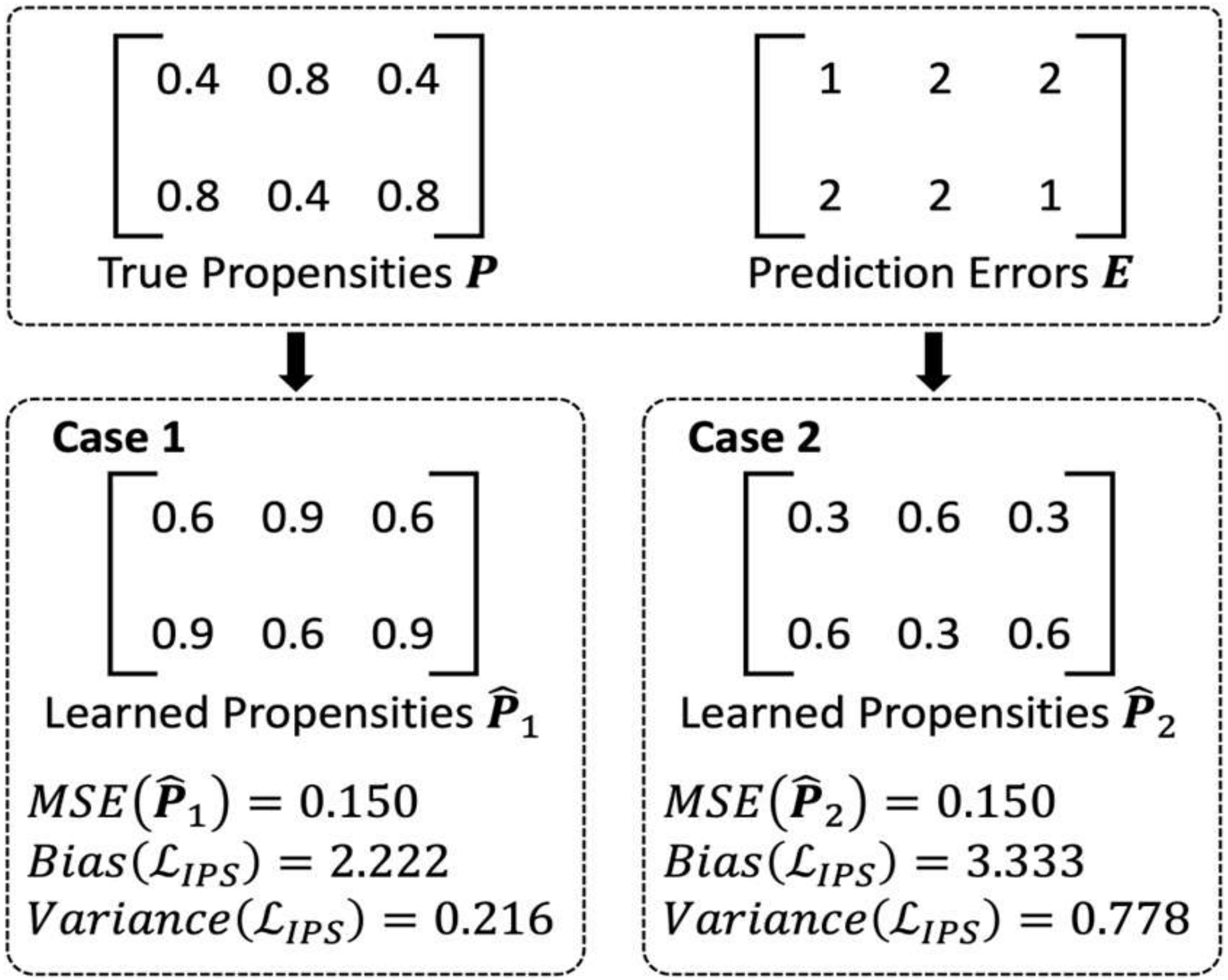
Method	$Bias[\mathcal{L}_{Method}(\hat{R})]$	$Variance[\mathcal{L}_{Method}(\hat{R})]$
IPS	$\frac{1}{\mathcal{D}}\left\vert\sum_{(u,i)\in\mathcal{D}} \frac{p_{u,i} - \hat{p}_{u,i}}{\hat{p}_{u,i}} e_{u,i}\right\vert$	$\frac{1}{\vert\mathcal{D}\vert^2}\sum_{(u,i)\in\mathcal{D}} \frac{p_{u,i}(1 - p_{u,i})}{\hat{p}_{u,i}^2} e_{u,i}^2$
DR	$\frac{1}{\mathcal{D}}\left\vert\frac{p_{u,i} - \hat{p}_{u,i}}{\hat{p}_{u,i}}(e_{u,i} - \hat{e}_{u,i})\right\vert$	$\frac{1}{\vert\mathcal{D}\vert^2}\sum_{(u,i)\in\mathcal{D}} \frac{p_{u,i}(1 - p_{u,i})}{\hat{p}_{u,i}^2}(e_{u,i} - \hat{e}_{u,i})^2$
MRDR		
DRMSE		

The bias and the variance can be affected by $\hat{p}_{u,i}$ and $\hat{e}_{u,i}$.

Motivation

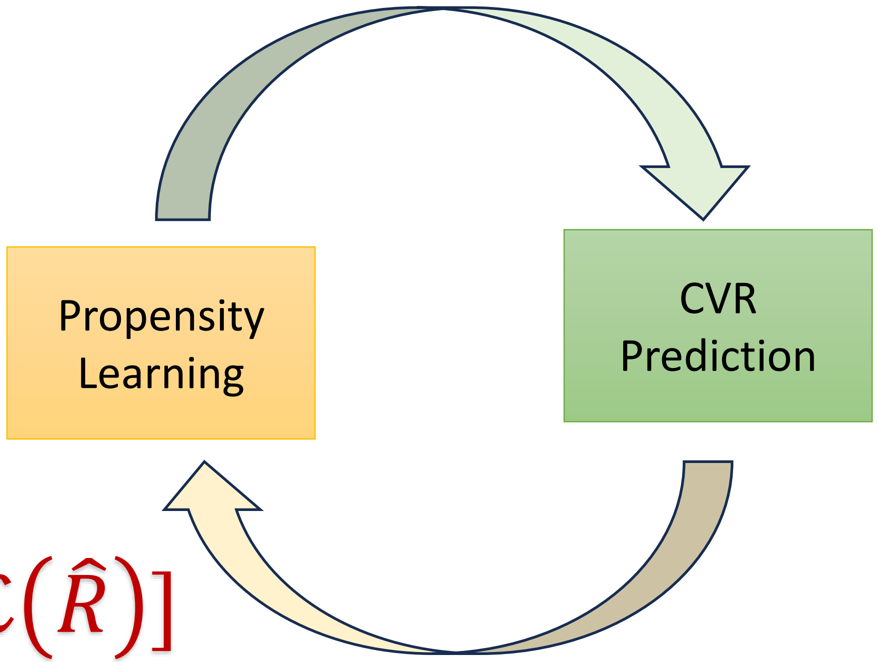
Previous works directly train the propensity learning model to fit the click data.

However, even if the model is trained on a dataset of unbiased click data, using its predicted propensity could still lead to high bias and variance of CVR estimators.



We propose a generalized propensity learning (GPL) framework, to directly reduce the bias and variance.

$\hat{p}_{u,i}$



$Bias[\mathcal{L}(\hat{R})]$
$Var[\mathcal{L}(\hat{R})]$

- Take the bias and variance of the CVR estimator into consideration when learning propensity.
- Train the propensity learning model and the CVR prediction model jointly.

Involve the bias and variance of CVR estimator into the loss function of propensity learning

Involve the bias and variance of CVR estimator into the loss function of propensity learning

- IPS

Upper bound of IPS bias: $\mathcal{L}_{IPS}^{Bias} = \frac{1}{\mathcal{D}} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} - 2o_{u,i}\hat{p}_{u,i} + \hat{p}_{u,i}^2}{\hat{p}_{u,i}^2} e_{u,i}^2$

Involve the bias and variance of CVR estimator into the loss function of propensity learning

- IPS

Upper bound of IPS bias: $\mathcal{L}_{IPS}^{Bias} = \frac{1}{\mathcal{D}} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} - 2o_{u,i}\hat{p}_{u,i} + \hat{p}_{u,i}^2}{\hat{p}_{u,i}^2} e_{u,i}^2$

Upper bound of IPS variance : $\mathcal{L}_{IPS}^{Var} = \frac{1}{|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}^2} e_{u,i}^2$

Involve the bias and variance of CVR estimator into the loss function of propensity learning

- IPS

Upper bound of IPS bias: $\mathcal{L}_{IPS}^{Bias} = \frac{1}{\mathcal{D}} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} - 2o_{u,i}\hat{p}_{u,i} + \hat{p}_{u,i}^2}{\hat{p}_{u,i}^2} e_{u,i}^2$

Upper bound of IPS variance : $\mathcal{L}_{IPS}^{Var} = \frac{1}{|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}^2} e_{u,i}^2$

$$\mathbf{min}\, \mathcal{L}_{IPS-GPL}^{Propensity} = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} CrossEntropy\left(o_{u,i}, \hat{p}_{u,i}\right) + \lambda[\beta \mathcal{L}_{IPS}^{Bias} + (1-\beta)\mathcal{L}_{IPS}^{Var}]$$

Involve the bias and variance of CVR estimator into the loss function of propensity learning

- DR-based (DR, MRDR, DRMSE)

Upper bound of DR bias:

$$\mathcal{L}_{DR-based}^{Bias} = \frac{1}{\mathcal{D}} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i} - 2o_{u,i}\hat{p}_{u,i} + \hat{p}_{u,i}^2}{\hat{p}_{u,i}^2} (e_{u,i} - \hat{e}_{u,i})^2$$

Upper bound of DR variance:

$$\mathcal{L}_{DR-based}^{Var} = \frac{1}{|\mathcal{D}|^2} \sum_{(u,i)\in\mathcal{D}} \frac{o_{u,i}}{\hat{p}_{u,i}^2} (e_{u,i} - \hat{e}_{u,i})^2$$

$$\mathbf{min}\, \mathcal{L}_{DR-GPL}^{Propensity} = \frac{1}{|\mathcal{D}|} \sum_{(u,i)\in\mathcal{D}} CrossEntropy(o_{u,i}, \hat{p}_{u,i}) + \lambda[\beta \mathcal{L}_{DR-based}^{Bias} + (1-\beta)\mathcal{L}_{DR-based}^{Var}]$$

Experiments & Results

- We take IPS, DR, MRDR, and DRMSE as our baseline models.

Dataset	Models	DCG@2	DCG@4	DCG@6
Yahoo	IPS	0.4895 ± 0.0033	0.6724 ± 0.0021	0.7999 ± 0.0017
	IPS-GPL	0.5600 ± 0.0061	0.7534 ± 0.0071	0.8730 ± 0.0053
		14.39%	12.04%	9.14%
	DR	0.5143 ± 0.0147	0.7132 ± 0.0129	0.8431 ± 0.0116
	DR-GPL	0.5519 ± 0.0052	0.7415 ± 0.0039	0.8626 ± 0.0024
		7.31%	3.96%	2.30%
	MRDR	0.5495 ± 0.0163	0.7445 ± 0.0175	0.8676 ± 0.0160
	MRDR-GPL	0.5634 ± 0.0026	0.7597 ± 0.0033	0.8825 ± 0.0020
		2.52%	2.04%	1.70%
	DRMSE	0.5462 ± 0.0125	0.7383 ± 0.0122	0.8600 ± 0.0120
	DRMSE-GPL	0.5678 ± 0.0038	0.7636 ± 0.0028	0.8841 ± 0.0015
		3.95%	3.42%	2.80%

Experiments & Results

- We apply GPL to IPS, DR, MRDR, and DRMSE, respectively.

Dataset	Models	DCG@2	DCG@4	DCG@6
Yahoo	IPS	0.4895 ± 0.0033	0.6724 ± 0.0021	0.7999 ± 0.0017
	IPS-GPL	0.5600 ± 0.0061	0.7534 ± 0.0071	0.8730 ± 0.0053
		14.39%	12.04%	9.14%
	DR	0.5143 ± 0.0147	0.7132 ± 0.0129	0.8431 ± 0.0116
	DR-GPL	0.5519 ± 0.0052	0.7415 ± 0.0039	0.8626 ± 0.0024
		7.31%	3.96%	2.30%
	MRDR	0.5495 ± 0.0163	0.7445 ± 0.0175	0.8676 ± 0.0160
	MRDR-GPL	0.5634 ± 0.0026	0.7597 ± 0.0033	0.8825 ± 0.0020
		2.52%	2.04%	1.70%
	DRMSE	0.5462 ± 0.0125	0.7383 ± 0.0122	0.8600 ± 0.0120
	DRMSE-GPL	0.5678 ± 0.0038	0.7636 ± 0.0028	0.8841 ± 0.0015
		3.95%	3.42%	2.80%

Experiments & Results

- The performance of IPS, DR, MRDR, and DRMSE are all improved by GPL.

Dataset	Models	DCG@2	DCG@4	DCG@6
Yahoo	IPS	0.4895 ± 0.0033	0.6724 ± 0.0021	0.7999 ± 0.0017
	IPS-GPL	0.5600 ± 0.0061	0.7534 ± 0.0071	0.8730 ± 0.0053
		14.39%	12.04%	9.14%
	DR	0.5143 ± 0.0147	0.7132 ± 0.0129	0.8431 ± 0.0116
	DR-GPL	0.5519 ± 0.0052	0.7415 ± 0.0039	0.8626 ± 0.0024
		7.31%	3.96%	2.30%
	MRDR	0.5495 ± 0.0163	0.7445 ± 0.0175	0.8676 ± 0.0160
	MRDR-GPL	0.5634 ± 0.0026	0.7597 ± 0.0033	0.8825 ± 0.0020
		2.52%	2.04%	1.70%
	DRMSE	0.5462 ± 0.0125	0.7383 ± 0.0122	0.8600 ± 0.0120
	DRMSE-GPL	0.5678 ± 0.0038	0.7636 ± 0.0028	0.8841 ± 0.0015
		3.95%	3.42%	2.80%

Experiments & Results

- More information can be found in our paper:
  - Demonstrate the improvement of performance on two real-world dataset (Coat Shopping and Yahoo!R3) and synthetic datasets
  - Show the reduction of bias and variance of CVR estimators on synthetic datasets
  - Hyper-parameter Study

Conclusion

- Propose a generalized propensity learning framework to learn effective propensities that directly reduce the bias and variance

- Derive the upper bounds of the bias and variance of IPS and DR-based estimators, used as new components to be optimized in the loss function of the propensity learning model.

- Empirically demonstrate the significant improvement and the effectiveness of the bias and variance reduction of GPL by experiments on real-world and semi-synthetic datasets.

Thank you for listening!

Q & A