Deep Learning Meets Double Machine Learning:

Causal Inference with Large-Scale Combinatorial Experiments

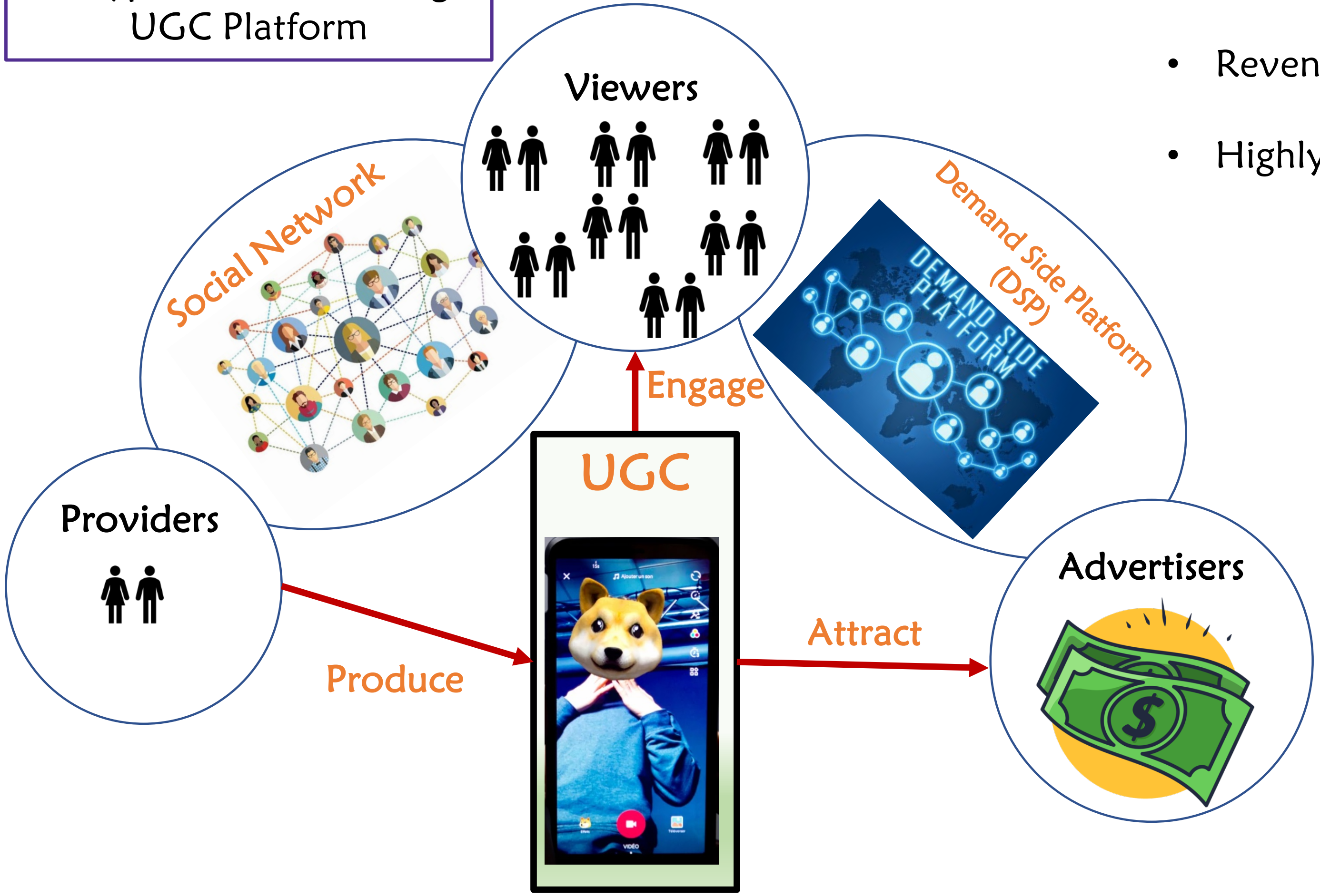
Renyu (Philip) Zhang

(Based on the joint work with Zikun Ye, Zhiqi Zhang, Dennis J. Zhang, Heng Zhang)

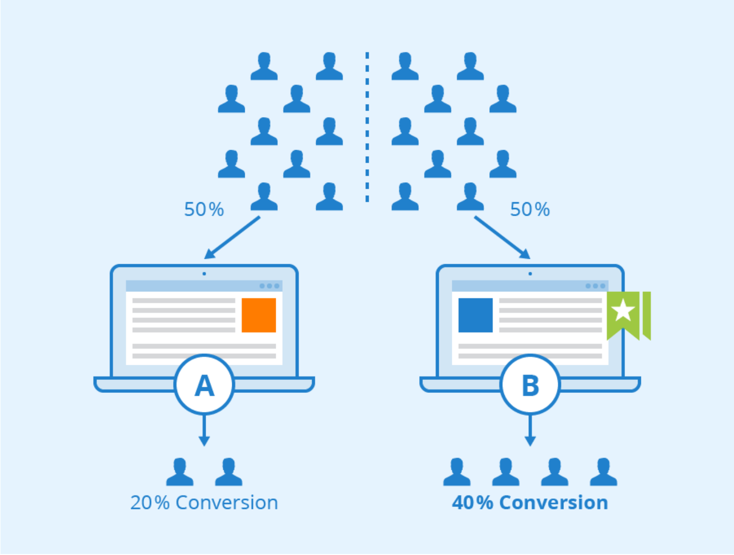
1. Introduction
2. Theory: Deep Learning, Double Machine Learning, and Asymptotics
3. Empirics: Implementation, Experiments, and Validations with Real and Synthetic Data

Video-Sharing UGC Platforms and A/B Tests

A Typical Video-Sharing UGC Platform

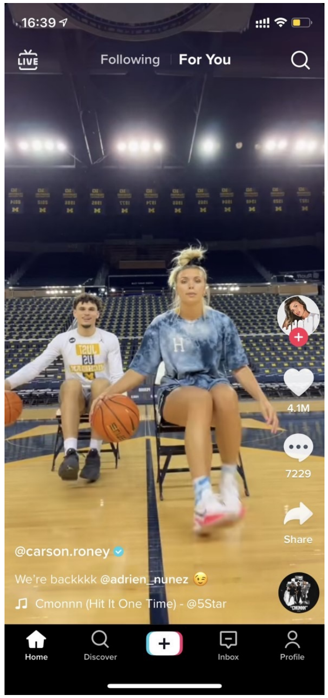


- MAU in billions (53.6% of world population).
- Revenue in hundreds of billions of USD per year.
- Highly individualized big data.

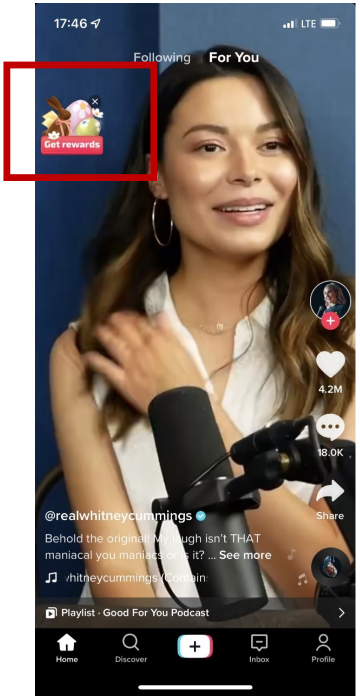


A/B Testing: The Decision Engine

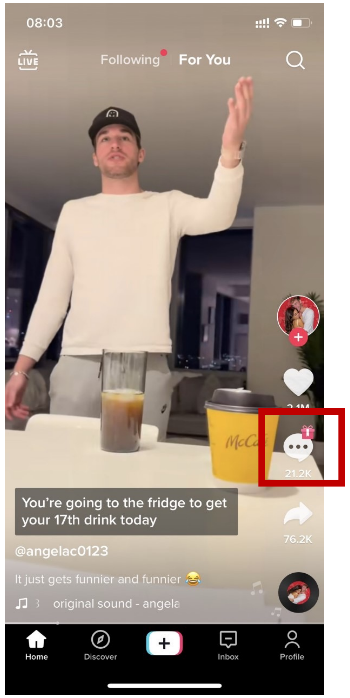
Multiple A/B Tests on Large-Scale Platforms



Baseline: Nothing



Treatment A:
“Get Rewards”

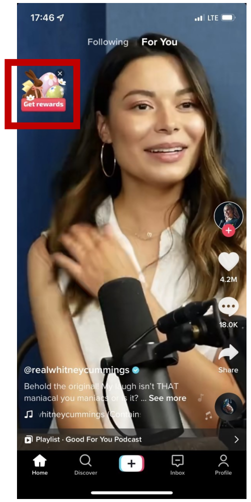

08:03
Following
For You
You're going to the fridge to get your 17th drink today
@angelac0123
It just gets funnier and funnier
original sound - angela
76.2K
Home
Discover
Inbox
Profile

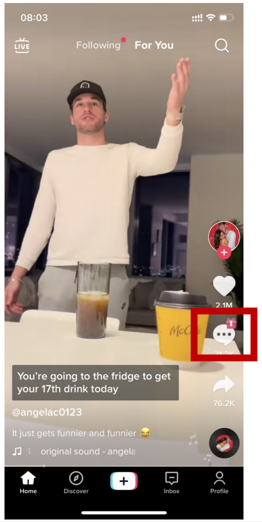
Treatment B:
“Send Gift”

- A large-scale platform launches hundreds of A/B tests everyday to fast iterate their operations and marketing strategies.
  - Usually under the orthogonal design.
- Users are independently treated by thousands of different A/B tests simultaneously.

How to estimate and infer the combined treatment effect of multiple A/B tests?

Solution 1: Linear Addition



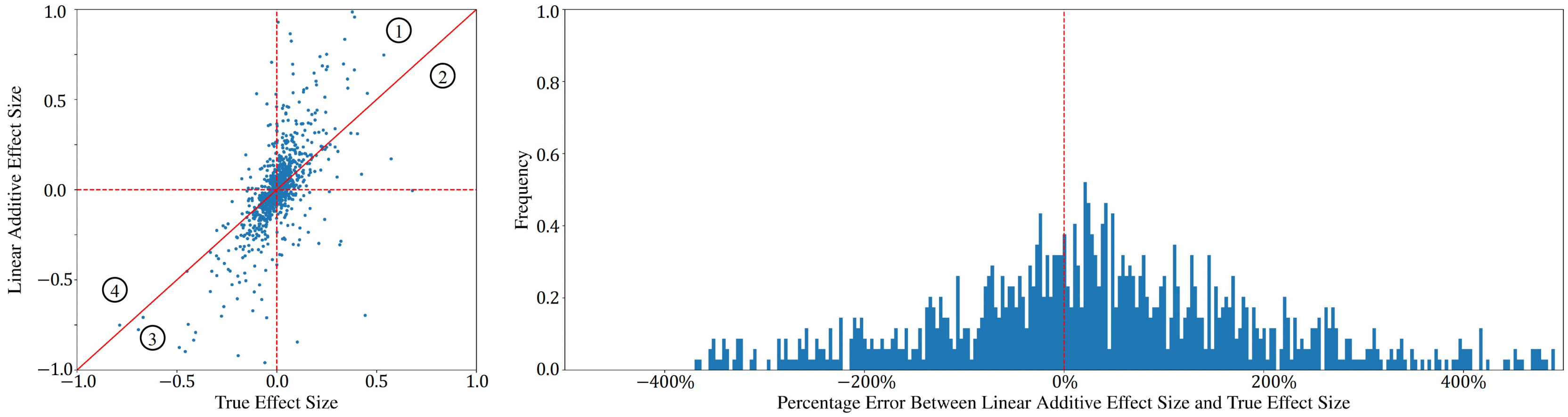


- Effect of "Get Rewards + Send Gift" = Effect of "Get Rewards" + Effect of "Send Gift"

Control	Treatment A	Treatment B
No button	Get Rewards	Send Gift

Limitations:

- Non-linearity: The effect of the combined treatment may not equal the sum of each.
    - Decreasing marginal return: (+7min) < (+3min) + (+5min)
    - Increasing marginal return: (+15min) > (+6min) + (+7min)

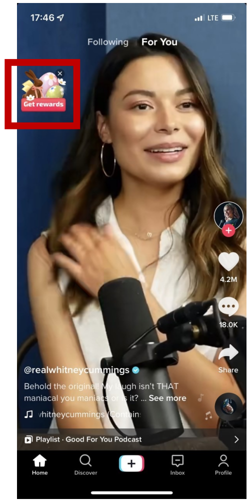


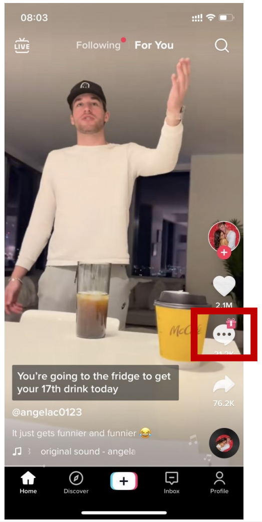
(a)

(b)

- Heterogeneity: The effect of the combined treatment may vary for different users.

Solution 2: Factorial Experiment





You're going to the fridge to get your 17th drink today

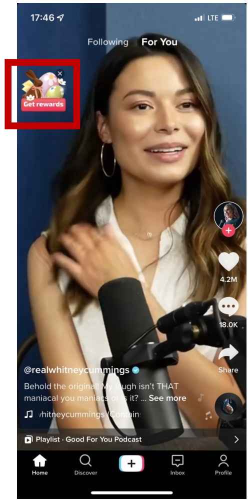
- Run an experiment with treatment combinations "Get Rewards", "Send Gift" and "Get Rewards + Send Gift".

Control	Treatment A	Treatment B	Treatment AB
No button	Get Rewards	Send Gift	Get Rewards and Send Gift

Limitation:

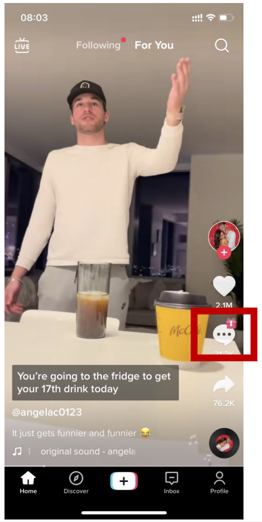
- m interventions generate 2^m treatment combinations.
- It is impossible to even assign only 1 user to each single treatment combination if $m > 30$.

Solution 3: End-to-End Deep Learning



Behold the original! My laugh isn't THAT maniacal you maniacs or is it? ... See more

♫ whitneycummings (Contains



You're going to the fridge to get your 17th drink today

It just gets funnier and funnier 😂

- Directly predict the outcome of each user under each treatment combination using end-to-end (e2e) deep learning (DL).

Control	Treatment A	Treatment B
No button	Get Rewards	Send Gift

Limitations:

- With unobserved treatment combinations, we cannot do causal inference with e2e DL (or any other pure machine learning methods such as uplift modeling).
  - Hard to obtain any economic and managerial insights.

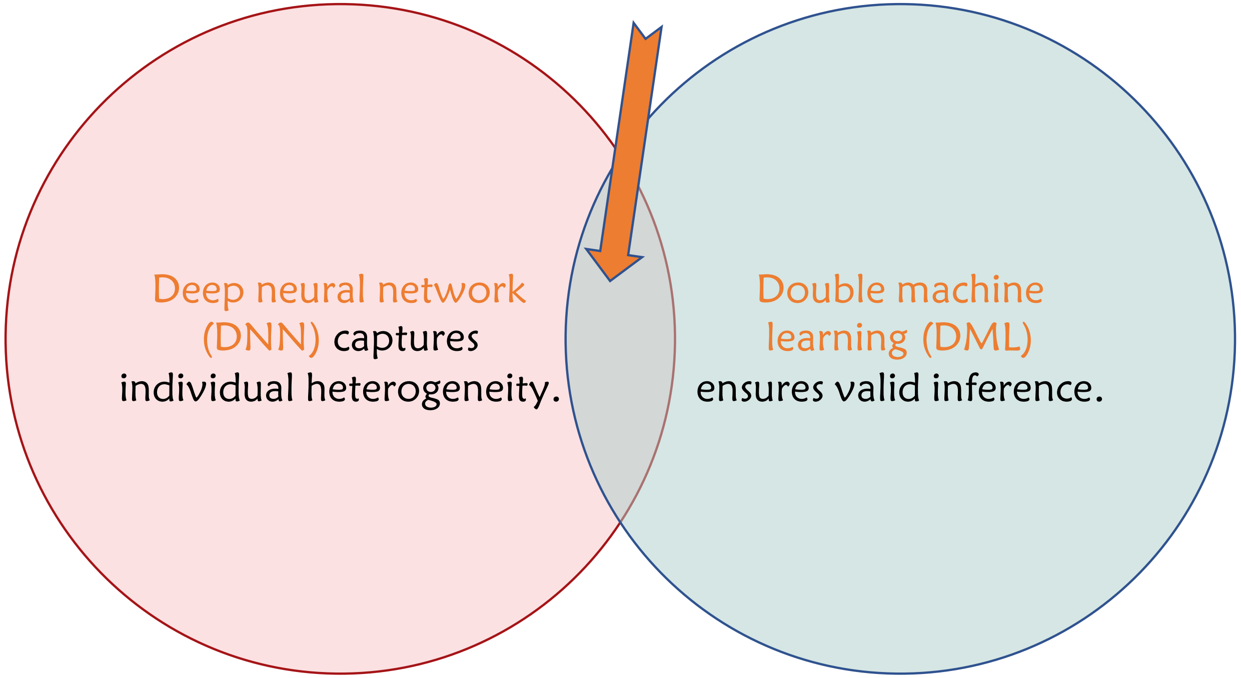
- How about the generalized random (causal) forests (Athey et al. 2019)?
  - Given the unobservable treatment combinations, causal trees/forests are essentially (locally) linear.

Key Research (and Business) Questions

Only observing the outcomes of a subset of treatment combinations:

- How to estimate and infer the effect of any treatment combination (i.e., ATE) under multiple A/B tests on the platform?

- How to identify the optimal treatment combination (i.e., best-arm identification)?



Deep neural network (DNN) captures individual heterogeneity.

Double machine learning (DML) ensures valid inference.

Related Literature

- Double/de-biased machine learning (DML): Correct the bias of a plug-in estimator through Neyman-orthogonal score functions.
  - Newey (1994), Chernozhukov et al. (2018, 2022), Farrell et al. (2020, 2021), Athey et al. (2018), Ellickson et al. (2022), Fan et al. (2022), etc.

- Valid estimation and inference under experimentation: Variance reduction and de-biasing.
  - Azevedo et al. (2020), Dasgupta et al. (2015), Athey et al. (2021), Johari et al. (2021), Bojinov et al. (2021), Candogan et al. (2021), Xiong et al. (2022), etc.

- Experiments on online platforms: Evaluating and optimizing the strategies of a large-scale online platform.
  - Ye et al. (2022), Zeng et al. (2022), Zhang et al. (2020), Cui et al. (2019, 2020), Feldman et al. (2021), Schwartz et al. (2017), etc.

Highlight of Main Contributions

- Theory

  - A new DL+DML framework.
  - Theoretical validity (consistency and normality) via Neyman orthogonality.

- Empirics

  - Implementation for real large-scale A/B tests on a video-sharing platform.
  - Better performance than the linear and DL benchmarks in ATE estimation and best-arm identification.

- Practice

  - Practical validations of DL+DML with data from large-scale field experiments (N>2,000,000).
  - Inspirations for future researchers and practitioners to apply DL+DML.

1. Introduction

2. Theory: Deep Learning, Double Machine Learning, and Asymptotics

3. Empirics: Implementation, Experiments, and Validations with Real and Synthetic Data

Deep Learning Framework: Setup

- The platform runs m A/B tests, each with a binary treatment. Use $\boldsymbol{t} \in \{0,1\}^m$ to denote a treatment combination.
    - For example, $m = 4, \boldsymbol{t} = (0, 0, 1, 0)'$ represents that the user is in the control condition of A/B test A, B, and D, and in the treatment condition of A/B test C.

- Outcome $Y \in \mathbb{R}$ and feature $\boldsymbol{X} \in \mathbb{R}^{d_X}$.

- Assume the data generating process (DGP):

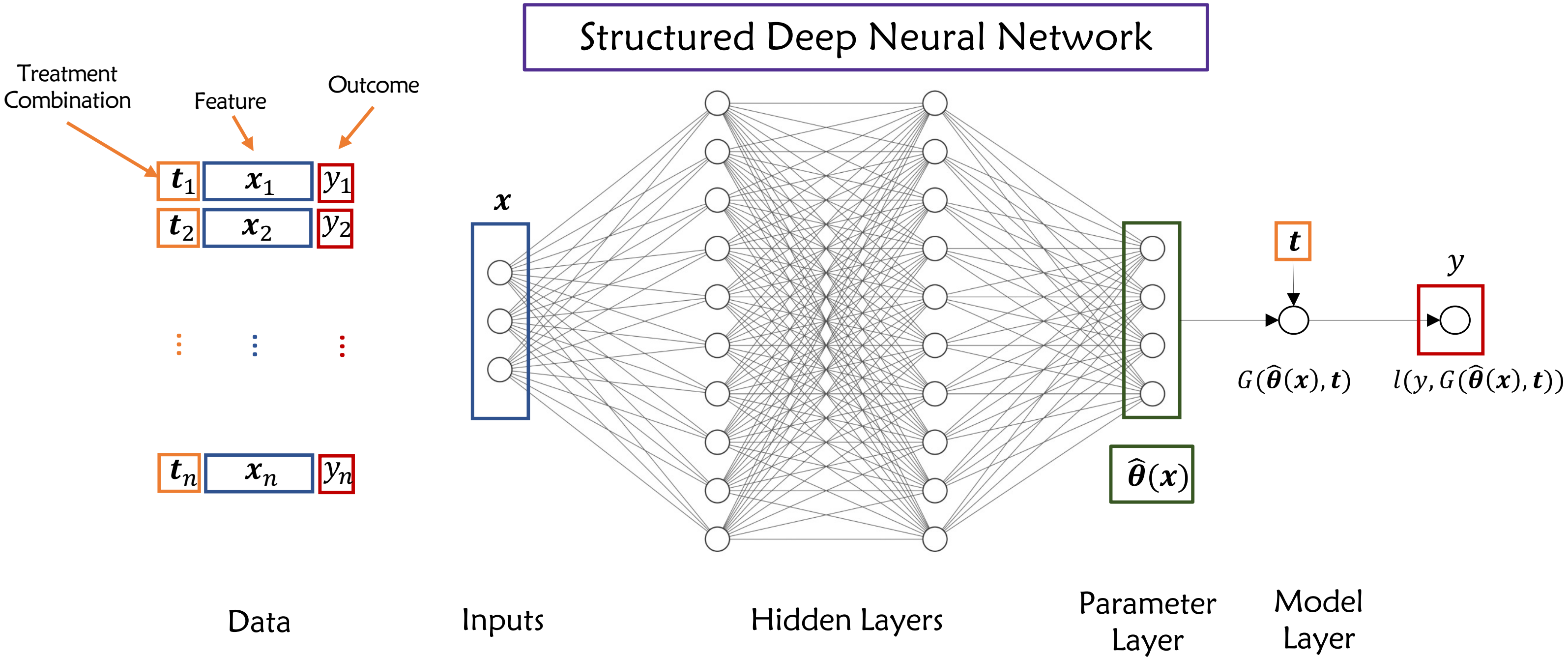
$$\mathbb{E}[Y|\boldsymbol{X} = \boldsymbol{x}, \boldsymbol{T} = \boldsymbol{t}] = G(\boldsymbol{\theta}^*(\boldsymbol{x}), \boldsymbol{t})$$

    - $G(\boldsymbol{\theta}, \boldsymbol{t})$ is the link function with a known (parametric) structure, mapping $\mathbb{R}^{d_X} \times \{0,1\}^m \rightarrow \mathbb{R}$
    - $\boldsymbol{\theta}^*(\cdot)$ is the (true) nonparametric function capturing HTE and obtained by $\boldsymbol{\theta}^*(\cdot) = \arg\min_{\theta \in \Theta} \mathbb{E}[l(Y, G(\boldsymbol{\theta}(\boldsymbol{X}), \boldsymbol{T}))]$, where $l(.,.)$ is the loss function (squared error).

- The parameters we are interested in estimating and inferring:
    - Average treatment effect (ATE): $\mu(\boldsymbol{t}) = \mathbb{E}[H(\boldsymbol{X}, \boldsymbol{\theta}^*(\boldsymbol{X}); \boldsymbol{t})] = \mathbb{E}[G(\boldsymbol{\theta}^*(\boldsymbol{X}), \boldsymbol{t}) - G(\boldsymbol{\theta}^*(\boldsymbol{X}), \boldsymbol{t}_o)]$, for all $\boldsymbol{t} \in \{0,1\}^m$, where $\boldsymbol{t}_o = (0,0, \dots, 0)'$.
    - Best-arm identification: $\boldsymbol{t}^* = \arg\max_{\boldsymbol{t} \in \{0,1\}^m} \mu(\boldsymbol{t})$.

- Two-stage procedure: (a) training; (b) estimation & inference.

Structured Deep Neural Nets



- Empirical estimator of $\boldsymbol{\theta}^*(.)$:

$$\widehat{\boldsymbol{\theta}}(.) = \underset{\boldsymbol{\theta}\in\mathcal{F}_{DNN}}{\arg\min} \frac{1}{n}\sum_{i=1}^{n} l(y_i, G(\boldsymbol{\theta}(\boldsymbol{x}_i), \boldsymbol{t}_i)),$$

which is obtained by SGD or Adam.

Note: The DL architecture is inspired by Farrell et al. (2020).

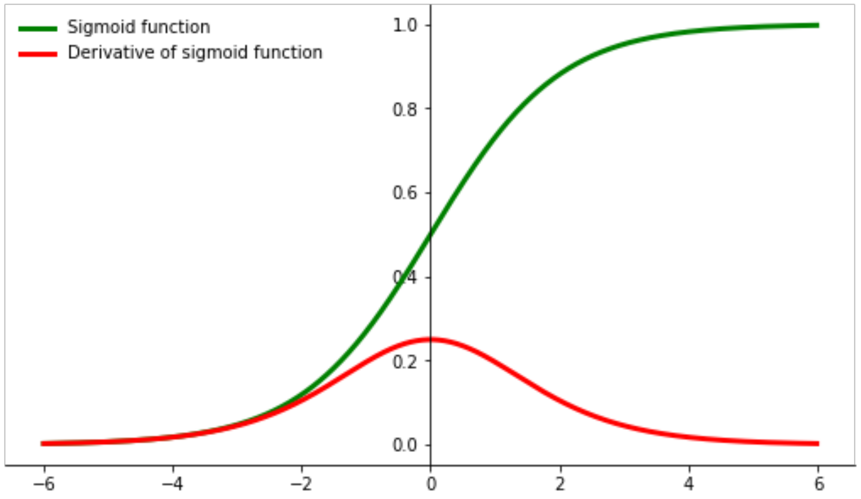
Link Function and Convergence

- We adopt the following link function $G(.,.)$ to approximate the true DGP:
  - Generalized Sigmoid Function:

$$G(\boldsymbol{\theta}^*(\boldsymbol{x}), \boldsymbol{t}) = \frac{\theta^*_{m+1}(\boldsymbol{x})}{1 + \exp(-(\theta^*_0(\boldsymbol{x}) + \theta^*_1(\boldsymbol{x})t_1 + \cdots + \theta^*_m(\boldsymbol{x})t_m))}$$

  - $\boldsymbol{\theta}(\boldsymbol{x})'\boldsymbol{t}$: The HTE of treatment combination $\boldsymbol{t}$ with respect to different $\boldsymbol{x}$.
  - The generalized sigmoid function captures both diminishing marginal return and/or increasing marginal return, and any possible ranges of potential outcomes (by $\theta^*_{m+1}(.)$).

Theorem. Under some regularity and network size assumptions on $\mathcal{F}_{DNN}$ and the treatment assignment mechanism (with m+2 observable combinations) of the A/B tests, $\widehat{\boldsymbol{\theta}}$ converges to $\boldsymbol{\theta}^*$ sufficiently fast $o(n^{-1/4})$ for inference (with subsequent debias).



$$\underbrace{\|\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}^*_k\|^2_{L_2(\boldsymbol{X})}}_{L_2\text{- Norm}} \leq C\left\{n^{-\frac{p}{p+d_{\boldsymbol{X}}}}\log^8 n + \frac{\log\log n}{n}\right\}$$

$$\underbrace{\mathbb{E}_n\left[(\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}^*_k)^2\right]}_{\text{Sample Average}} \leq C\left\{n^{-\frac{p}{p+d_{\boldsymbol{X}}}}\log^8 n + \frac{\log\log n}{n}\right\}$$

- p: Smoothness of the DNN class.

Debias with Neyman Orthogonal Score

- The plug-in (PI) estimator for ATE:

$$\hat{\mu}_{PI}(\boldsymbol{t}) = \frac{1}{n}\sum_{i=1}^{n} H\big(\boldsymbol{x}_i, \widehat{\boldsymbol{\theta}}(\boldsymbol{x}_i); \boldsymbol{t}\big) = \frac{1}{n}\sum_{i=1}^{n}\left[G\big(\widehat{\boldsymbol{\theta}}(\boldsymbol{x}_i), \boldsymbol{t}\big) - G\big(\widehat{\boldsymbol{\theta}}(\boldsymbol{x}_i), \boldsymbol{t}_o\big)\right]$$

- A critical issue with the PI estimator: Insufficient convergence speed to the true ATE (we need root-N consistency).
  - Additional biases and inconsistencies from perturbations of $\widehat{\boldsymbol{\theta}}(.)$ because of regularization and/or the variations in $\boldsymbol{X}$.

- Solution: Neyman Orthogonal Score.
  - Moment conditions: $\mathbb{E}[\psi(\boldsymbol{W}, \mu, \boldsymbol{\theta}^*)] = 0$ (ψ is the score function, $\boldsymbol{W} = (Y, (\boldsymbol{X}, \boldsymbol{T})')'$ is the data, μ is the ATE, and $\boldsymbol{\theta}^*$ is the true parameter).
  - Neyman Orthogonality: $\partial_{\boldsymbol{\theta}}\mathbb{E}[\psi(\boldsymbol{W}, \mu, \boldsymbol{\theta})]|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*} = 0$.
  - Under Neyman orthogonality, even though $\widehat{\boldsymbol{\theta}}$ slightly perturbs from the true value $\boldsymbol{\theta}^*$, it does not affect the moment conditions.
    - The bias of $\widehat{\boldsymbol{\theta}}$ will not affect the moment conditions, so it will not significantly change the subsequent estimator $\hat{\mu}$.

Theorem. Under nonrestrictive regularity assumptions, $\psi(\boldsymbol{w}, \boldsymbol{\theta}, \boldsymbol{\Lambda}; \boldsymbol{t}) - \mu(\boldsymbol{t})$ is a Neyman Orthogonal score, where

$$\underbrace{\psi(\boldsymbol{w}, \boldsymbol{\theta}, \boldsymbol{\Lambda}; \boldsymbol{t})}_{\text{Influence Function}} = \underbrace{H(\boldsymbol{x}, \theta(\boldsymbol{x}); \boldsymbol{t})}_{\text{Plug-In Estimator}} - \underbrace{\partial_{\boldsymbol{\theta}} H(\boldsymbol{x}, \boldsymbol{\theta}(\boldsymbol{x}); \boldsymbol{t})\boldsymbol{\Lambda}(\boldsymbol{x})^{-1}\partial_{\boldsymbol{\theta}} l\big(y, G(\boldsymbol{\theta}(\boldsymbol{x}), \boldsymbol{t})\big)}_{\text{De-bias Term}}, \text{ with } \boldsymbol{\Lambda}(\boldsymbol{x}) = \mathbb{E}[\partial^2_{\boldsymbol{\theta}} l(y, G(\boldsymbol{\theta}(\boldsymbol{X}), \boldsymbol{t}))|\boldsymbol{X} = \boldsymbol{x}].$$

- $\boldsymbol{t}$: In the data.
- $\boldsymbol{t}$: Want to estimate.

- Remark: The influence function is derived based on the pathwise derivative approach in semi-parametric statistics (Newy 1994, Chernozhukov et al. 2018, Farrell et al. 2020).

Cross-Fitting and Asymptotic Normality

- To avoid over-fitting, we apply cross-fitting:
  - The training set is split into S non-overlapping subsets $S_1, S_2 \cdots S_S$. $\widehat{\boldsymbol{\theta}}_{\boldsymbol{s}}$ is trained on S_s^c, the complement of S_s.

$$\psi(\boldsymbol{w}, \boldsymbol{\theta}, \boldsymbol{\Lambda}; \boldsymbol{t}) = H(\boldsymbol{x}, \theta(\boldsymbol{x}); \boldsymbol{t}) - \partial_{\boldsymbol{\theta}} H(\boldsymbol{x}, \boldsymbol{\theta}(\boldsymbol{x}); \boldsymbol{t}) \boldsymbol{\Lambda}(\boldsymbol{x})^{-1} \partial_{\boldsymbol{\theta}} l\big(y, G(\boldsymbol{\theta}(\boldsymbol{x}), \boldsymbol{t})\big)$$

$$\hat{\mu}(\boldsymbol{t}) = \frac{1}{S} \sum_{i=1}^{S} \hat{\mu}_s(\boldsymbol{t}), \qquad \hat{\mu}_s(\boldsymbol{t}) = \frac{1}{|S_s|} \sum_{j \in S_s} \psi(\boldsymbol{w}_j, \widehat{\boldsymbol{\theta}}_s(\boldsymbol{x}_j), \widehat{\boldsymbol{\Lambda}}_s(\boldsymbol{x}_j); \boldsymbol{t})$$

$$\widehat{\Psi}(\boldsymbol{t}) = \frac{1}{S} \sum_{i=1}^{S} \widehat{\Psi}_s(\boldsymbol{t}), \qquad \widehat{\Psi}_s(\boldsymbol{t}) = \frac{1}{|S_s|} \sum_{j \in S_s} (\psi(\boldsymbol{w}_j, \widehat{\boldsymbol{\theta}}_s(\boldsymbol{x}_j), \widehat{\boldsymbol{\Lambda}}_s(\boldsymbol{x}_j); \boldsymbol{t}) - \hat{\mu}(\boldsymbol{t}))^2$$

Theorem. Under nonrestrictive regularity assumptions,

$$\sqrt{n/\widehat{\Psi}(\boldsymbol{t})}(\hat{\mu}(\boldsymbol{t}) - \mu(\boldsymbol{t})) \rightarrow_d \mathcal{N}(0,1)$$

- ATE Estimator: $\hat{\mu}(\boldsymbol{t})$.
- $(1-\alpha)$-Confidence Interval: $[\hat{\mu}(\boldsymbol{t}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{\Psi}(\boldsymbol{t})}{n}}, \hat{\mu}(\boldsymbol{t}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\widehat{\Psi}(\boldsymbol{t})}{n}}]$.
- Partial observability: $\boldsymbol{t}$ can be an unobservable treatment combination.

We call the entire framework as Debiased Deep Learning (DeDL).

Best-Arm Identification

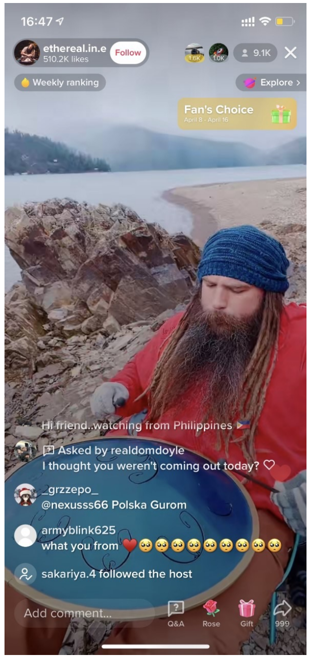
- The true best-arm: $\boldsymbol{t}^* = \arg\max_{\boldsymbol{t}\in\{0,1\}^m} \mu(\boldsymbol{t})$; the estimated best-arm: $\hat{\boldsymbol{t}}^* = \arg\max_{\boldsymbol{t}\in\{0,1\}^m} \hat{\mu}(\boldsymbol{t})$.
- The advantage of $\hat{\boldsymbol{t}}^*$ over $\boldsymbol{t}$: $\tau(\boldsymbol{t}) \coloneqq \mu(\hat{\boldsymbol{t}}^*) - \mu(\boldsymbol{t})$; the estimator for $\tau(\boldsymbol{t})$: $\hat{\tau}(\boldsymbol{t}) = \hat{\mu}(\hat{\boldsymbol{t}}^*) - \hat{\mu}(\boldsymbol{t})$.
- The influence function for $\tau(\boldsymbol{t})$: $\psi(\boldsymbol{w}, \boldsymbol{\theta}, \boldsymbol{\Lambda}; \hat{\boldsymbol{t}}^*) - \psi(\boldsymbol{w}, \boldsymbol{\theta}, \boldsymbol{\Lambda}; \boldsymbol{t})$, via which the SE of $\hat{\tau}(\boldsymbol{t})$ can be derived.

Theorem. Under nonrestrictive regularity assumptions, $\hat{\tau}(\boldsymbol{t})$ is a consistent estimator of $\tau(\boldsymbol{t})$, and $\sqrt{n}\,(\hat{\tau}(\boldsymbol{t}) - \tau(\boldsymbol{t}))$ converges to a normal distribution.

- To verify $\hat{\boldsymbol{t}}^* = \boldsymbol{t}^*$, it suffices to do one-sided tests for the Hypotheses $\tau(\boldsymbol{t}) > 0$, where $\boldsymbol{t} \in \{0,1\}^m$.
- The DeDL framework can be applied to estimating and inferring a wide rage of quantities of interest, with the influence function properly (re-)derived. Examples:
  - ATE of a personalized policy to adopt (estimated) optimal treatment combination $\hat{\boldsymbol{t}}^*$ for each user.
  - Policy evaluation for any personalized policy π that maps a user feature $\boldsymbol{x}$ to a distribution on the treatment space $\{0,1\}^m$.

1. Introduction
2. Theory: Deep Learning, Double Machine Learning, and Asymptotics
3. Empirics: Implementation, Experiments, and Validations with Real and Synthetic Data

Field Setting



Hi friend..watching from Philippines

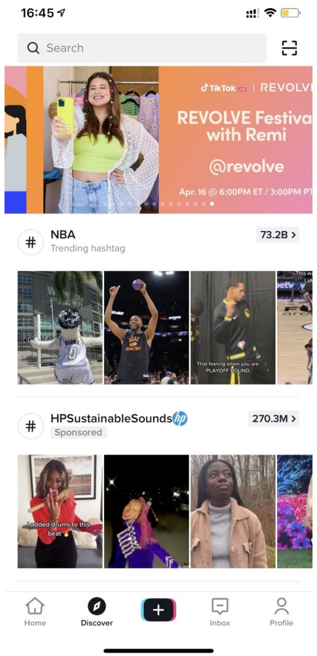
Asked by realdomdoyle
I thought you weren't coming out today?

grzzepo
@nexusss66 Polska Gurom

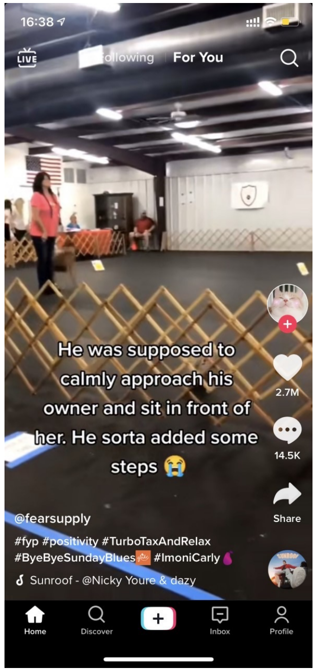
armyblink625
what you from ❤️🥺🥺🥺🥺🥺🥺🥺🥺🥺

sakariya.4 followed the host

Live Page



Discover Page



He was supposed to calmly approach his owner and sit in front of her. He sorta added some steps 😭

@fearsupply

#fyp #positivity #TurboTaxAndRelax #ByeByeSundayBlues #ImoniCarly

Sunroof - @Nicky Youre & dazy

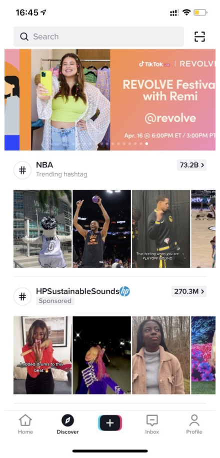
For You Page

- A Chinese online short-video sharing platform (referred to as Platform O hereafter).
- 350 million+ DAU, half-billion+ MAU, 20 million+ USD advertising revenue per day.
- Platform O launches hundreds of A/B tests everyday to fast iterate their business operations.
- We consider $m = 3$ major A/B tests on the algorithmic upgrades of the 3 pages on the left.

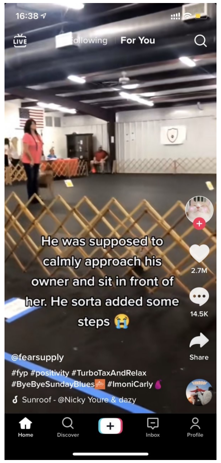
Objective: (a) Estimate and infer ATE; (b) Best-arm identification.

A/B Tests, Data, and Ground-Truth

Live Page



Discover Page


16:38
LIVE
Following
For You
He was supposed to calmly approach his owner and sit in front of her. He sorta added some steps 😭
@fearsupply
#fyp #positivity #TurboTaxAndRelax #ByeByeSundayBlues #ImoniCarly
Sunroof - @Nicky Youre & dazy
2.7M
14.5K
Share
Home
Discover
Inbox
Profile

For You Page

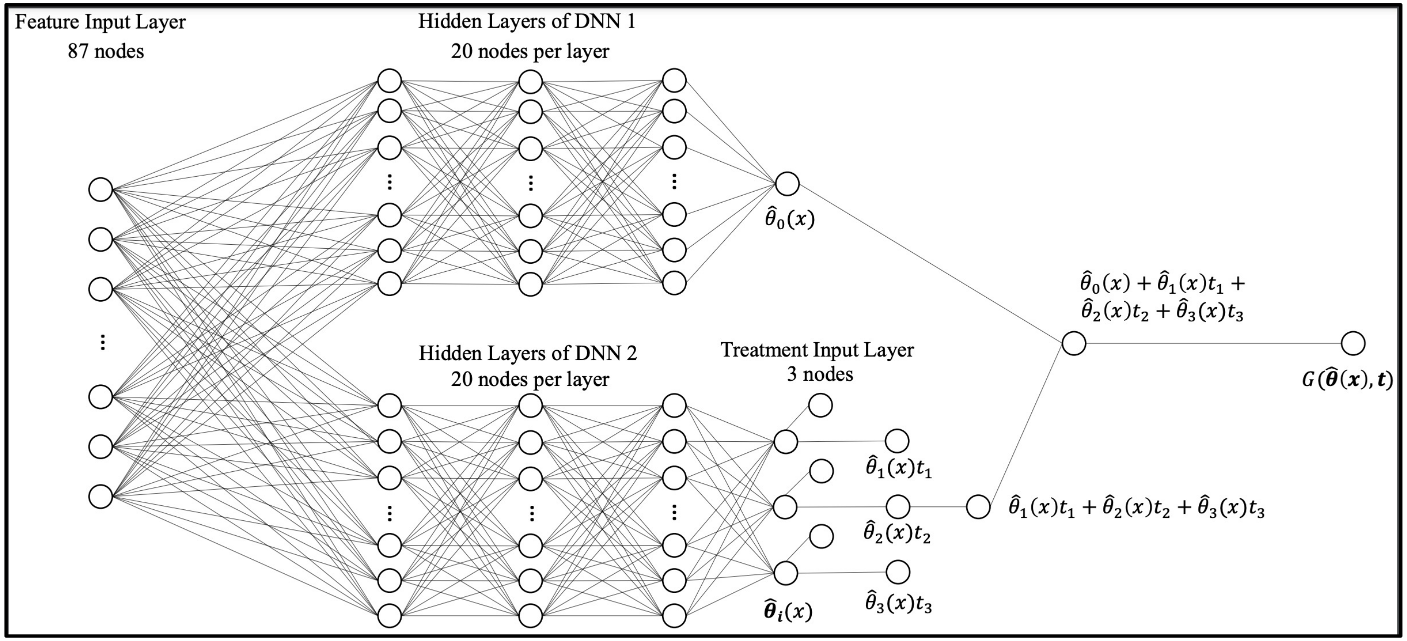
- Duration: Jan 10, 2021-Feb. 01, 2021.
- Sample size: 2,066,606 (roughly 258,325 under each $\boldsymbol{t} \in \{0,1\}^3$)
- Y = Total video-watching time of a user per day.
- $\boldsymbol{X}$ = User demographics (e.g., gender) and pre-treatment behaviors (e.g., the number of active days 1 week before the experiment).
- Randomization checks are passed, so users under different treatment combinations are comparable.

Treatment Combination (t)	Ground-Truth ATE (Scaled)	Observable?	Number of Users
(0,0,0)	0.000%	Observable	258,249
(0,0,1)	1.091%**	Observable	258,340
(0,1,0)	-0.267%	Observable	258,367
(1,0,0)	0. 758%*	Observable	258,321
(1,1,1)	2.121%****	Observable	258,375
(1,1,0)	0.689%	Unobservable	258,480
(1,0,1)	2.299%****	Unobservable	258,305
(0,1,1)	1.387%***	Unobservable	258,172

Note:
- Observable means observable for the estimators.
- The relative ATEs are reported to protect sensitive data.
- True best-arm: $\boldsymbol{t}^* = (1,0,1)$
- *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001.

Implementation of the DeDL Framework

- DGP: $\mathbb{E}[Y|\boldsymbol{X}=\boldsymbol{x},\boldsymbol{T}=\boldsymbol{t}] = G(\boldsymbol{\theta}^*(\boldsymbol{x}),\boldsymbol{t}) = \frac{\theta_4^*(\boldsymbol{x})}{1+exp\left(-(\theta_0^*(\boldsymbol{x})+\theta_1^*(\boldsymbol{x})t_1+\theta_2^*(\boldsymbol{x})t_2+\theta_3^*(\boldsymbol{x})t_3)\right)}$



Feature Input Layer
87 nodes
Hidden Layers of DNN 1
20 nodes per layer
$\hat{\theta}_0(x)$
$\hat{\theta}_0(x)+\hat{\theta}_1(x)t_1+$
$\hat{\theta}_2(x)t_2+\hat{\theta}_3(x)t_3$
$G(\hat{\boldsymbol{\theta}}(\boldsymbol{x}),\boldsymbol{t})$
Hidden Layers of DNN 2
20 nodes per layer
Treatment Input Layer
3 nodes
$\hat{\theta}_1(x)t_1$
$\hat{\theta}_1(x)t_1+\hat{\theta}_2(x)t_2+\hat{\theta}_3(x)t_3$
$\hat{\theta}_2(x)t_2$
$\hat{\boldsymbol{\theta}}_i(x)$
$\hat{\theta}_3(x)t_3$

- The DNNs are trained with data from the observable treatment combinations.

- One DNN for $\hat{\theta}_0$ (dropout rate=0.1) and the other for $(\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3)$ (dropout rate=0.2). Each has 3 hidden layers; each layer has 20 nodes. All use ReLU as the activation function.

- The third DNN for $\hat{\theta}_4$ is trained as a linear layer.

Benchmarks

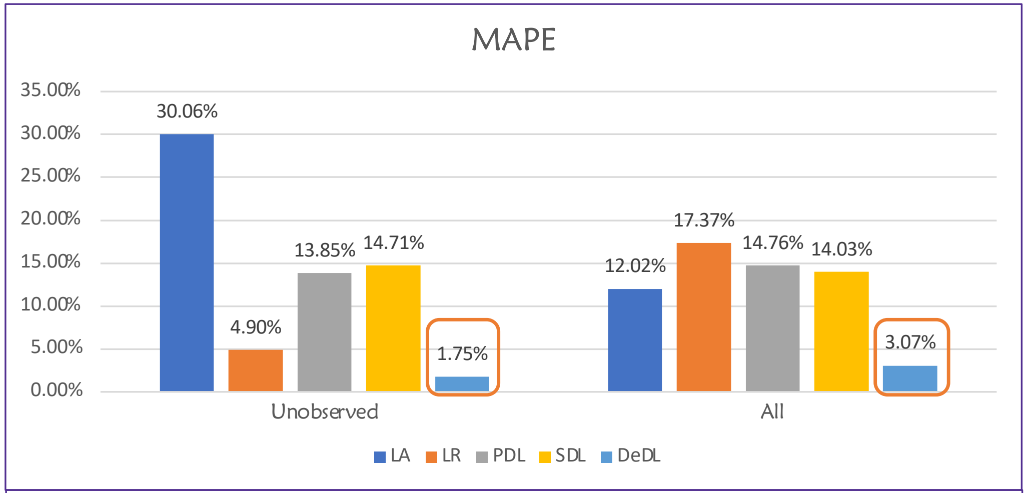
- Linear Addition (LA): Assume that the ATE of different individual treatments are linearly and independently additive.
  - Effect of "Get Rewards + Send Gift" = Effect of "Get Rewards" + Effect of "Send Gift"

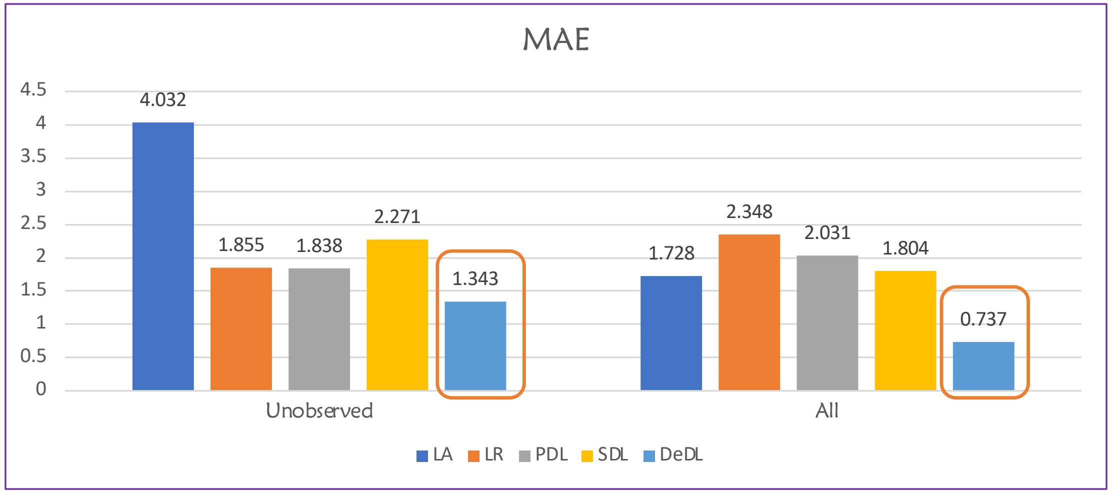
- Linear Regression (LR): Regress Y on $(T', X')'$ and predict the outcomes of unobservable treatment combinations by linear extrapolation.
  - Still a linear approach, but better leverages the user features.

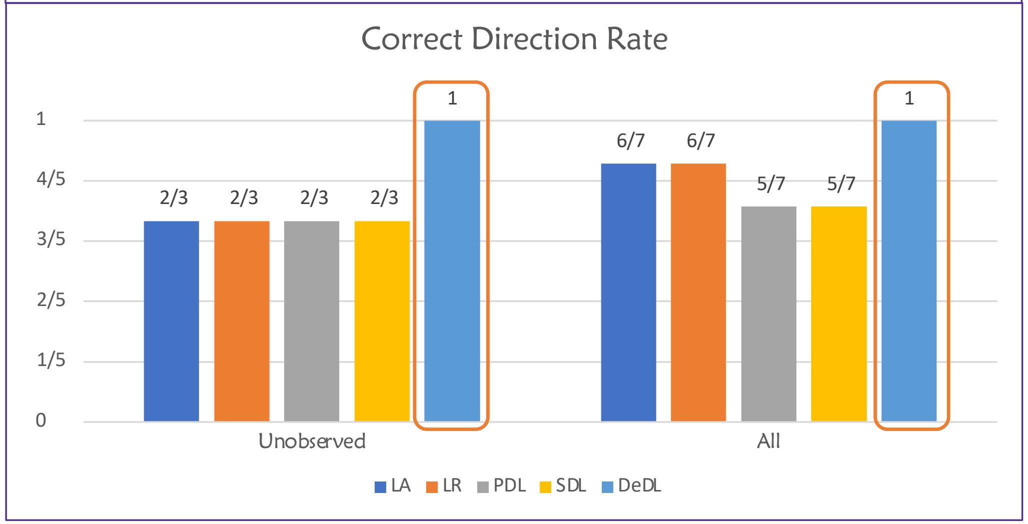
- Pure Deep Learning (PDL): Apply a generic DNN with $(T', X')'$ as the inputs to predict the outcomes of unobservable treatment combinations.
  - Fully leverages the predictive power of DNN but without valid inference.

- Structured Deep Learning (SDL): Apply the same DNN as DeDL without debias to predict the outcomes of unobservable treatment combinations.
  - Comparing DeDL with SDL highlights the value of bias correction through DML.

ATE Estimation and Inference





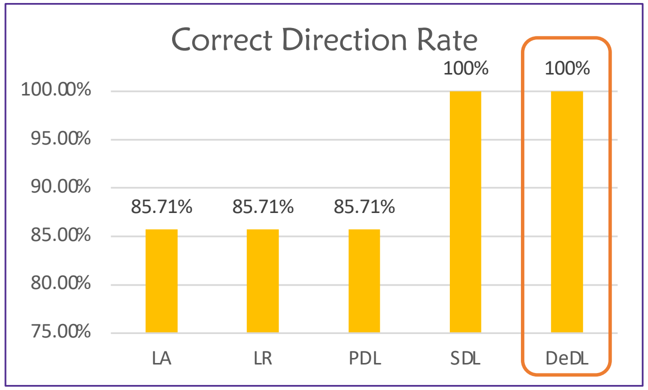


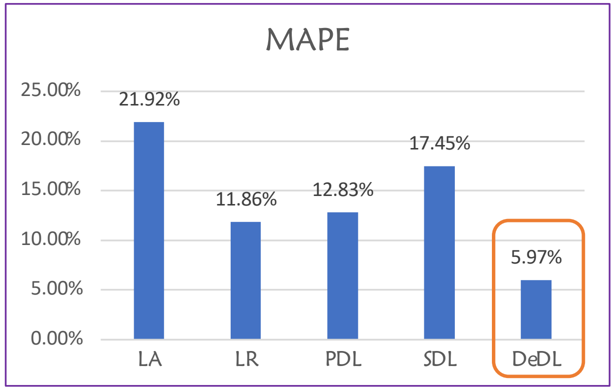
- The performance metrics are evaluated against the ground truth ATE with respect to 3 (resp. 7) unobservable (resp. all) treatment combinations.

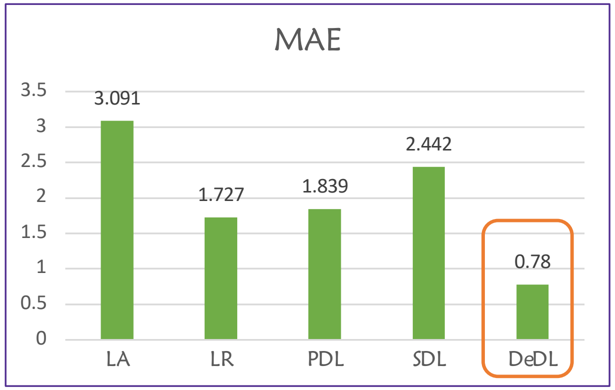
- Correct Direction = Correctly identifying the statistical significance and sign of ATE.

- Key insights:
  - The empirical results validate DeDL in a **field setting**!
  - Naive application of DNNs does NOT outperform linear benchmarks.
  - Bias-correction via Neyman orthogonality substantially improves the performance of DNNs for every treatment combination.

Best-Arm Identification





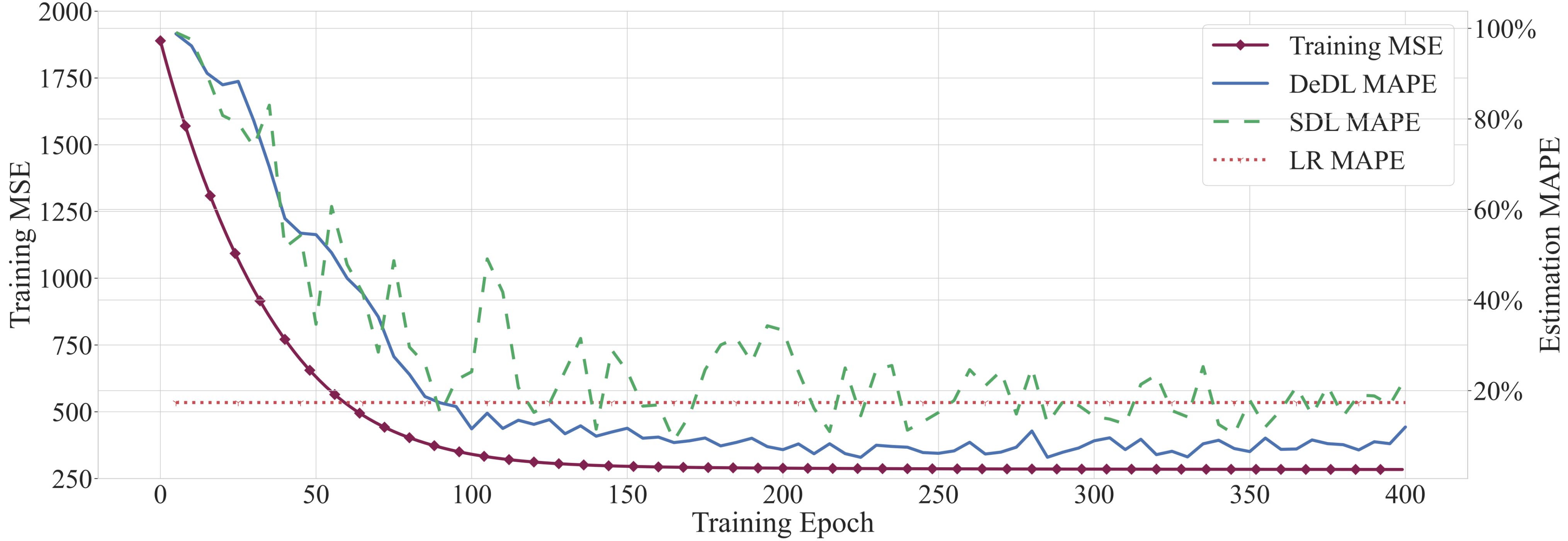


Note:

- We report the CDR, MAPE and MAE of estimating $\tau(\boldsymbol{t})$ against the ground-truth for LA, LR, PDL, SDL, and DeDL.

- DeDL and SDL can reliably identify the optimal treatment combination, $\hat{\boldsymbol{t}}^* = \boldsymbol{t}^* = (1,0,1)$.
- DeDL outperforms the benchmarks for better inferring the advantage of $\hat{\boldsymbol{t}}^*$ over other treatment combinations.
  - $\tau(\boldsymbol{t})$'s are more accurately predicted by DeDL.

From Training to Inference



- If the DNN is not designed and/or not well-trained, the ATE estimation via DeDL will have a terrible performance (MAPE>60%).
- If the DNN performs well, DeDL will consistently beat linear and DL benchmarks without debiasing.
- The DNN training error serves as an important indicator for the quality of second-stage estimation leveraging debiasing.

Insights from Synthetic Data

Good News

- The advantage of DeDL expands when the number of A/B tests m is larger.
- If the link function $G(.,.)$ is correctly specified, DeDL performs well even when additional biases are introduced in the training procedure.

Bad News

- If the link function $G(.,.)$ is seriously misspecified, DeDL may perform poorly.
  - Vulnerability under model misspecification.
  - Model misspecification can be detected by DNN training error.
  - Recipe: (i) abandon the debias term; (ii) auto-debias (Chernozhukov et al. 2022).

Takeaways

- DeDL framework: A new DL+DML framework to estimate and infer the causal effects of multiple A/B tests on large-scale platforms with unobservable outcomes.
  - Theoretical valid for inference via Neyman orthogonality.

- Implementation: Real large-scale A/B tests (N>2,000,000) on Platform O.
  - Better performance than the linear and DL benchmarks in ATE estimation and best-arm identification.
- Practice: Inspirations for future researchers and practitioners to apply DL+DML in other important settings for program evaluation with experimental or observational data.

Code: https://github.com/zikunye2/deep_learning_based_causal_inference_for_combinatorial_experiments