An End-to-end Learning Approach for Counterfactual Generation and Individual Treatment Effect Estimation

Feilong Wu, Kejing Yin, William K. Cheung
Department of Computer Science, *Hong Kong Baptist University*
Hong Kong, China
{csflwu, cskjyin, william}@comp.hkbu.edu.hk

***Abstract*—Estimating the causal effect due to an intervention is important for many applications, such as healthcare. Unobserved counterfactuals make unbiased treatment effect estimation non-trivial. Among existing approaches, counterfactual generation which augments observational data with generated pseudo counterfactuals has been found promising for reducing the bias. These methods typically take a two-stage approach for the counterfactual generation and treatment effect estimation. Therefore, the counterfactual generation could be sub-optimal. To this end, we propose to jointly optimize the auxiliary models for generating the counterfactuals and the outcome estimation models. In particular, we demonstrate the viability by first connecting a counterfactual outcome generator with a reparameterized VAE model, and then learning them in an end-to-end fashion using the EM algorithm. Our evaluation results based on synthetic and semi-synthetic datasets show that a simple causal effect VAE model learned together with the counterfactual outcome generator can outperform a number of SOTA models for treatment effect estimation[1].**

***Index Terms*—treatment effect estimation, counterfactual generation, causal inference**

I. Introduction

Estimating the causal effect of an intervention is a fundamental problem across diverse domains. Applications include predicting a patient's health status after a therapy or the economic growth of a society after economic stimulus. Apart from understanding the effect of an intervention on a population level, there is much interest in developing methods to estimate the effect at the individual level, which can be crucial for life-critical areas such as healthcare [1], [2], [3], [4].

Leveraging available observational data for individualized treatment effect (ITE) estimation has been widely explored, as this can avoid the need to collect data from costly or sometimes infeasible randomised experiments. Yet unbiased estimation of treatment effect from observational data is non-trivial where issues like selection bias and confounding factors should be carefully handled. The former one refers to the bias in selecting the subjects and thus the observational data. The latter refers to the factors affecting both the treatment and the outcome, which in turn may also induce selection bias when the set of variables that predispose selection into the treatment are also related to the outcome.

To address the unbiased estimation challenge, the counterfactual framework has been commonly explored. Once we observe the outcome of applying a treatment to a particular subject, we will not be able to observe the outcome if the treatment is not applied (a.k.a. counterfactuals). As treatments are typically given to the subjects based on the anticipated outcome, selection bias becomes unavoidable. Various approaches have been proposed in the literature to address the challenge, including inverse propensity weighting [5], [6], [7], confusion regularization [8], [9], data augmentation [10], [11], [12], etc. Among them, the data augmentation approach has been shown effective in some recent works [2], [11], [13], [10], [12]. For example, GANITE [10] generates counterfactuals to augments with observed data to be fed to an ITE generator for learning the model in a GAN setting. Meta-learning methods [11], [12], [13] make use of the pseudo treatment effects obtained using the models learned for the treated and the control separately to train the second-stage models. Yet, the aforementioned methods are based on the two-stage design and thus lacks the mechanism to guide the generation of pseudo counterfactuals to achieve better model estimation in the second stage of training.

In addition, some confounding factors can be hidden and further complicate the unbiased estimation. Deep generative models have been explored to alleviate the bias, including the representative causal effect variational autoencoder (CEVAE) [14] and its variants (e.g., TEDVAE [15]). They infer latent confounders through the observed proxy variables, requiring only weak assumptions on the nature of the confounders and the data generation process. Despite their promising results, how to learn these VAE-based models from the observational data with unobserved counterfactuals properly considered remains open.

To this end, we explore the possibility to train the pseudo counterfactual generator and the ITE estimation model end-to-end with the objective to reduce the bias. In particular, we extend the CEVAE with a refined model parameterization and a pseudo counterfactual generator, and train them end-to-end. Our proposed model named *CEMVAE* explicitly represents the observed factuals and unobserved counterfactuals where the

[1]The source code for the implementation of our proposed method can be found on GitHub (https://github.com/FeilongWu/Unbiased-Treatment-Effect-Estimation/tree/main/CEMVAE)

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00040

unobserved counterfactuals are taken as missing information. The model is then learned using the EM algorithm so that the unobserved counterfactuals can be estimated via the iterations of the E-step and the M-step.

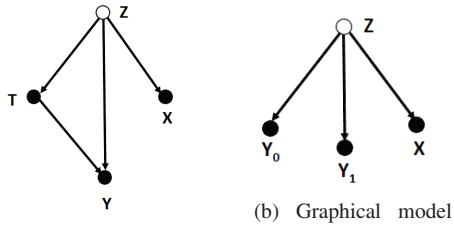
To demonstrate the effectiveness of CEMVAE, we apply it to both synthetic and semi-synthetic (healthcare) data and compare its performance with some state-of-the-art ITE estimation methods which also adopt the data augmentation approach. Our results obtained show that the proposed CEMVAE which merely adopts the basic causal effect VAE model together with a learnable counterfactual generator can outperform most of the SOTA estimation methods.

II. Related Work

In this section, we provide an overview of methods proposed for addressing the selection bias challenge as presented in Section I. They can be categorized into three mainstream approaches: i) *inverse propensity weighting* (IPW) which tries to balance the objective loss by weighting its terms with the inverse of the propensity score [16], which refers to a conditional density that treatment is assigned to a subject given its features (covariates) and is constant for any covariates from an unbiased dataset, so that the resulting objective will be the same to the loss calculated using data from a randomized control trial [5], [6], [7]; ii) *confusion regularization* which manipulates the feature projections via regularization so that the treatment group representations in the projected space are indistinguishable from those of the control group [8], [9]; and iii) *data augmentation* where the training data is augmented with pseudo counterfactual outcomes/treatment effect to achieve deconfounding [10], [11], [12]. In general, these approaches, in the first stage, use auxiliary models to generate fixed pseudo counterfactual outcomes/treatment effect. In the second stage, a main estimator is trained based on the augmented dataset. Our approach differentiates from those in that its pseudo counterfactual outcomes are adaptively improved along with the main estimator through end-to-end training.

Regarding treatment effect estimation, various models have been proposed in the literature. Causal trees [17], [2] are used to estimate treatment effect by dividing the feature space into partitions, each representing a path to a leaf node indicating the treatment effect. The partitions can be constructed in a recursive and greedy manner. Alternatively, deep generative models like CEVAE [14] have been proposed when latent variables were introduced as proxies of the underlying confounding factors. TEDVAE [15] introduces additional latent variables to capture also instrumental and risk factors to achieve a better unbiased estimation. GANITE [10] learns the ITE estimation model using the generative adversarial network (GAN) setting with the objective to achieve better generalization performance.

Our proposed method (CEMVAE) is similar to the methods like GANITE [10] in the sense that we leverage pseudo counterfactual generation for balanced training as the supervised loss regarding outcomes due to control (e.g., no treatment)



(a) Causal Model (b) Graphical model adopted by CEMVAE

Fig. 1: The white nodes refer to a hidden variable and black nodes indicate observed variables. (a) The causal model for our problem setting. (b) The graphical model for our proposed method based on the augmented dataset.

can be computed for treatment group and vice versa. The performance then depends on the quality of the generated pseudo counterfactuals w.r.t the ground truth.

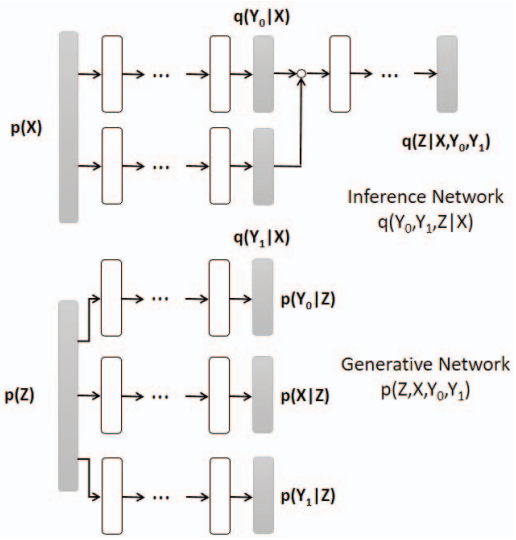


Fig. 2: Both generative and inference networks constitute the overall architecture of CEMVAE. The white boxes represent the parametrized deterministic neural network transitions, while the gray boxes correspond to drawing samples from the respective distribution, and the white circle represents the concatenation of data.

III. Preliminaries

Let $X \in \mathbb{R}^{D_x}$ denote the feature vector of a subject with a size D_x, for which x denotes its realization. Let $T \in \{0, 1\}$ denote the binary variable indicating the treatment given to a subject, and t the realized value. $T = 1$ implies treatment is given to the subject, and $T = 0$ indicates no treatment (control). With a realized value of y, $Y \in \mathbb{R}$ denotes the outcome variable of a subject to some treatment. We consider a set of observations $\mathcal{D}_{obs} = \{x_i, t_i, y_i\}_{i=1}^{n}$, where each factual treatment is observed with a propensity score $\pi(t_i|x_i)$,

a conditional density [16]. Since the outcomes due to $T = 0$ and $T = 1$ can be generated by distinct processes, we let the variable $Y_0 \in \mathbb{R}$ denote the potential outcome due to $T = 0$ and variable $Y_1 \in \mathbb{R}$ denote the potential outcome due to $T = 1$. Their realized values are respectively denoted as y_0 and y_1. The conditional average treatment effect (CATE) is defined as $\tau(X) := \mathbb{E}[Y|X, do(T = 1)] - \mathbb{E}[Y|X, do(T = 0)]$. In an observational dataset, only the factual outcome y_t is observed and the counterfactual outcome y_{1-t} is hidden.

In this paper, we consider the same causal graph as in [14] (Fig. 1a), where the variable Z is introduced as the latent confounder for confounding both T and Y. Also, similar to [14], [15], we assume the following:

Assumption 1 (Overlap)**.** *Every subject has a non-zero probability to receive a treatment, i.e.,* $0 < P(T = 1|Z = z) < 1, \forall z \in Z$.

This is a common assumption to ensure that the potential outcome can be identified for both $T = 0$ and $T = 1$.

IV. Methodology

In this section, we first provide an overview of the key ideas behind the proposed CEMVAE model where the data augmentation approach is incorporated into a VAE-based model for unbiased treatment effect estimation. Then, we present the detailed deep model architecture for the implementation and the training algorithm.

A. Overview

The unavoidable selection bias in observational data can result in situations unfavorable for unbiased estimation. For instance, the size of the treatment group and that of the control group can be highly imbalanced in the observational data. Also, the population distribution of the treated $P(X|T = 1)$ and the control $P(X|T = 0)$ for training can be unrepresentative to the whole population $P(X)$ which could be encountered during testing. This is a result of $P(Z|T = 0) \neq P(Z|T = 1)$ when X becomes a proxy variable of the latent variable Z (see Fig. 1a). The selection bias can be reduced if $P(Z|T = 0) = P(Z|T = 1) = P(Z)$. Data augmentation methods typically make use of auxiliary models for generating counterfactual outcomes to alleviate the bias on estimating the potential outcomes (POs) or treatment effect [11], [10], [13], [12]. Most of the existing methods adopt the two-stage approach under which the augmented data is fixed after generation in the first stage, followed by learning the estimators in the second stage. The accuracy of the augmented data can affect the performance of the estimators to be learned.

In this paper, our conjecture is that instead of taking the two-stage approach, methodologies which can generate more relevant counterfactuals are expected to improve the overall treatment effect estimation. Particularly, one-stage algorithms can be explored to allow iterative improvement on the accuracy of the generated counterfactuals by jointly optimizing the auxiliary models and main estimators. To achieve that, we first make reference to the CEVAE where the selection bias is yet to be addressed as mentioned in [14]. We then reparameterize the corresponding causal model (Fig. 1a) to form a graphical model as shown in Fig. 1b. Different from the CEVAE, the reparameterized graphical model has separate variables to represent the outcomes with or without treatment. We show in Section IV-B that the treatment effect is still identifiable given the reparameterized graphical model.

Figure 2 shows the corresponding deep model architecture of our proposed implementation. In particular, a single network is learned to estimate the latent variable Z given the feature X. Then, the outcomes with or without treatment are estimated at the same time, where only one of the outcomes can be factual and the other is counterfactual. This is fundamentally different from models like the CEVAE. For the CEVAE, two networks, one is for the treated and one for the control, are learned separately via a switching mechanism. Such design makes the quality of the counterfactual outcome estimation hard to be considered during the model training. For our proposed model, as depicted in Figure 2, two auxiliary models are introduced in the inference network for generating counterfactual outcomes. To learn the auxiliary models and the potential outcome estimators together under an end-to-end learning framework, we adopt the EM algorithm. Details of the overall model and the learning algorithm are presented in Section IV-C.

B. Identifiability

The treatment effect has been shown identifiable in CEVAE [14] if $P(Z, X, T, Y)$ can be recovered. Similarly, the identifiability of conditional average treatment effect (CATE) can be achieved if $P(Z, X, Y_0, Y_1)$ is recovered by our method, which leverages on data augmentation (see Appendix A).

C. Learning With Counterfactuals

We denote the observed data by $\mathcal{D}_{obs} = \{x_i, t_i, y_i\}_{i=1}^n$ and the unobserved counterfactuals by $\mathcal{D}_{cf} = \{x_i, 1 - t_i, y_i\}_{i=1}^n$, where n is the dataset size. Each x_i in $\mathcal{D}_{cf}$ takes the same value as that in $\mathcal{D}_{obs}$ while the treatment is reversed (i.e., $1 - t_i$). The value of y_i in $\mathcal{D}_{cf}$ is missing and to be imputed. If the "complete" data with both observed data and the unobserved counterfactuals, that is $\mathcal{D}_{obs} \bigcup \mathcal{D}_{cf}$, could be available, we can achieve the propensity score $\pi(T = 1|X) = \pi(T = 0|X)$, which resonates an unbiased dataset. The graphical model for the VAE model becomes Fig. 1b.

As observing the potential outcome with or without treatment for each subject together is impossible, what we would like to achieve is to estimate the counterfactual one as far as possible via imputation. To this end, we propose to estimate the imputed values by maximizing the training objective that takes *expectation over counterfactuals* on the evidence lower bound (ELBO) of the observational data. Inspired by [18], an unsupervised missing data recovery framework learning representations of heterogeneous data, CEMVAE adopts the idea of missing data recovery and connects it with downstream supervised learning tasks for estimating potential outcomes in an end-to-end manner. As a result, the pseudo counterfactual labels can be adaptively improved.

We parameterize the causal model as follows. Similar to CEVAE [14], the details of the VAE factors adopted are:

$$p(z_i) = \prod_{j=1}^{D_z} \mathcal{N}(z_{i,j}|0,1) \qquad p(x_i|z_i) = \prod_{j=1}^{D_x} p(x_{i,j}|z_i)$$
$$p(y_{t_i,i}|z_i, T=t_i) = \mathcal{N}(\mu = \bar{\mu}_{t_i}, \sigma^2 = \bar{\sigma}^2_{t_i})$$

with D_z and D_x corresponding to the dimensions of z_i and x_i, for which i denotes the index of the i-th subject. We represent $p_\psi(x_i|z_i)$, $p_{\epsilon_0}(y_{0,i}|z_i, T=0)$, and $p_{\epsilon_1}(y_{1,i}|z_i, T=1)$ as the generative models parameterized by ψ, ϵ_0, and ϵ_1, respectively. The distribution parameters $\bar{\mu}_{(\cdot)}$ and $\bar{\sigma}^2_{(\cdot)}$ are calculated by the associated models. According to Fig. 1b, the posterior distribution of Z depends on both POs and X. Hence, we have the following posterior approximation:

$$q(z_i|x_i, y_{t_i,i}, y_{1-t_i,i}) = \prod_{j=1}^{D_z} \mathcal{N}(\mu_j = \hat{\mu}_{z,i,j}, \sigma_j^2 = \hat{\sigma}^2_{z,i,j}).$$

Again, we let θ represent the parameters of the inference model $q_\theta(z_i|x_i, y_{0,i}, y_{1,i})$. The parameters $\hat{\mu}_{(\cdot)}$ and $\hat{\sigma}_{(\cdot)}$ are estimated by the associated models. Note that only $y_{0,i}$ or $y_{1,i}$ is available in a given dataset during training.

To impute the counterfactual outcome, we need the following auxiliary models:

$$q_{\phi_T=t_i}(y_{t_i,i}|x_i, T=t_i) = \mathcal{N}(\mu = \hat{\mu}_{t_i,i},\ \sigma^2 = \hat{\sigma}^2_{t_i,i})$$

where $\phi_{T=0}$ and $\phi_{T=1}$ (or simply as ϕ_0 and ϕ_1) are model parameters. Since the ELBO used for training is derived from the given data, we can take the expectation over the counterfactual outcome [18] by employing the EM algorithm to recover the counterfactuals iteratively in an end-to-end fashion. We thus define the expected ELBO for training as:

$$\mathcal{L}'_{ELBO}(x, y_t) = \mathbb{E}_{y_{1-t}}[\mathcal{L}_{\theta,\psi,\epsilon_0,\epsilon_1,\phi_0,\phi_1}(x, y_t, y_{1-t})].$$

Since only the conditional counterfactual likelihood and the posterior latent variable likelihood depend on y_{1-t}, the above lower bound can be expanded as:

$$\begin{aligned}\mathcal{L}'_{ELBO}(x, y_t) = \\ \sum_{i=1}^{n} \mathbb{E}_z[\mathbb{E}_{y_{1-t_i}}[\log p_{\epsilon_{1-t_i},\phi_{1-t_i}}(y_{1-t_i,i}|z_i) - \\ \mathbb{KL}(q_\theta(z_i|x_i, y_{t_i,i}, y_{1-t_i,i})|p(z_i))]] \\ + \mathbb{E}_z[log\, p_{\epsilon_{t_i}}(y_{t,i}|z_i) + \log p_\psi(x_i|z_i)], \quad (1)\end{aligned}$$

where $\mathbb{KL}(\cdot)$ is the Kullback–Leibler divergence.

For the E-step, as direct calculation of the counterfactual outcomes is intractable, we employ the two aforementioned auxiliary models $q_{\phi_0}(y_{0,i}|x_i, T=0)$ and $q_{\phi_1}(y_{1,i}|x_i, T=1)$. During training, the counterfactual outcomes can be generated by sampling from the distribution modeled by $q_{\phi^*_{1-t_i}}(y_{1-t_i,i}|x_i, T=1-t_i)$ where $*$ indicates the parameters learned up to the current iteration.

For the M-step, Eq.1 is maximized. The samples from the factors $q_{\phi_0}(y_{0,i}|x_i, T=0)$, $q_{\phi_1}(y_{1,i}|x_i, T=1)$, and $q_\theta(z_i|x_i, y_{0,i}, y_{1,i})$ can be drawn using the reparametrization trick [19]. All the parameters $\{\theta, \psi, \epsilon_0, \epsilon_1, \phi_0, \phi_1\}$ can then be optimized. To ensure the accuracy of the auxiliary models, we further add the following auxiliary objective function:

$$\mathcal{L}_{aux} = \sum_{i=1}^{n} \log q_{\phi_{t_i}}(y_{t_i,i}|x_i, T=t_i), \quad (2)$$

and thus the overall objective function becomes:

$$\mathcal{L}_{CEMVAE} = \mathcal{L}'_{ELBO} + \alpha \mathcal{L}_{aux}, \quad (3)$$

where $\alpha \in \mathbb{R}^+$. We call our proposed model Counterfactual Expectation Maximization VAE (CEMVAE). Fig. 2 shows the overall deep model architecture. Note that both the auxiliary models and the VAE backbone may have different convergence rates, so the auxiliary models can be pre-trained for better stability (see pseudo code in Appendix B).

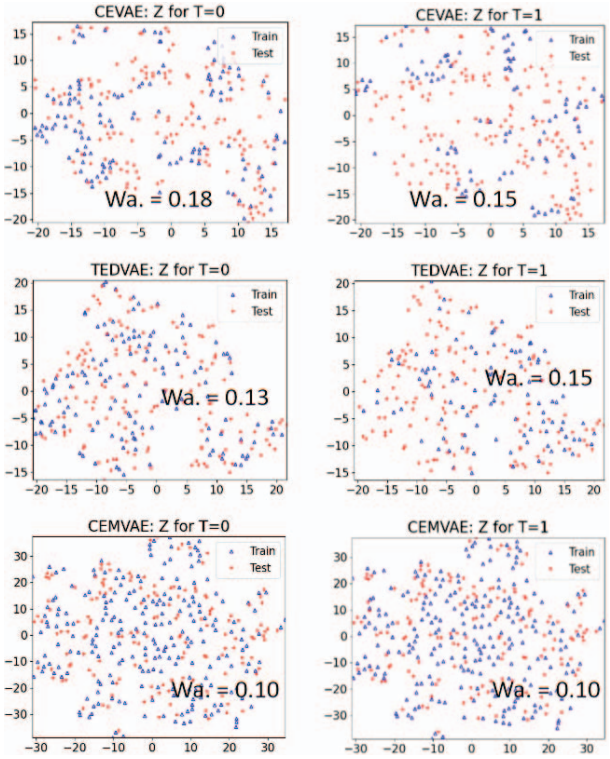


Fig. 3: This figure shows the t-SNE projections of the inferred $\mathbb{E}(Z^*)$ for the three VAE-based models during training and testing based on the synthetic dataset. The projection is done individually for each model. For CEVAE, Z refers to either Z_0 or Z_1 depending on the given actual treatment [14]. For TEDVAE [15], Z is a concatenation of Z_c and Z_y. To quantify the difference of Z distributions between training and test sets, the Wasserstein distance (Wa.) is calculated as reported in the sub-figures.

V. EXPERIMENTS

We evaluate the performance of the proposed method using a synthetic dataset where Z can be controlled, and two semi-synthetic datasets, IHDP [20] and eICU [21], where the X and T are real.

A. Datasets

a) Synthetic data generation: We generate a synthetic data according to causal graph shown in Fig. 1a, where the generation process for each sample is given as follows:

$$\begin{aligned}
\mathbf{z}_i &\sim \mathcal{N}(1, 0.5); \\
\mathbf{x}_{con,i}|\mathbf{z}_i &\sim \mathcal{N}(\mathbf{z}_i^3 \mathbf{M}_{con}, 0.01); \\
\mathbf{x}_{bin,i}|\mathbf{z}_i &\sim \mathrm{Bern}(\sigma(\mathbf{z}_i \mathbf{M}_{bin})); \\
t_i|\mathbf{z}_i &\sim \mathrm{Bern}(\sigma(\mathbf{z} \cdot \mathbf{v}_t - 1)); \\
y_i|\mathbf{z}_i &\sim \mathcal{N}(\mu_{t_i}, 0.02),
\end{aligned}$$

where each entry in $\mathbf{z}_i \in \mathbb{R}^8$ is sampled from $\mathcal{N}(1, 0.5)$, $\mathbf{x}_i$ is the feature vector with two sub-vectors $\mathbf{x}_{con,i}$ of continuous values and $\mathbf{x}_{bin,i}$ of binary values, $\mathbf{z}_i^3$ is an element-wise cubic vector of z_i, $\mathbf{M}_{con}$ and $\mathbf{M}_{bin}$ are a 8×10 matrix and a 8×3 matrix respectively, $\sigma(\cdot)$ denotes a logistic function, $\mathbb{1}$ is a vector of ones, $\mu_{t_i} = 0.8(\mathbf{z}_i - 0.5) \cdot (\mathbf{z}_i - 0.5) \cdot \mathbf{v}_y + \frac{2.2(t_i - 1)(\mathbb{1} \cdot \mathbf{z}_i)}{5} + 1$, and $\mathbf{v}_y \in \mathbb{R}^8$. The elements of M_{con}, M_{bin}, $\mathbf{v}_t$ as well as v_y are all sampled from $\mathcal{N}(0, 1)$, which are kept fixed during the data generation process. In synthetic data, CATE is a linear function of Z, while the nonlinear situation is simulated in the following datasets.

b) Semi-synthetic datasets: IHDP dataset is constructed based on the Infant Health and Development Program [20]. The confounders are the measurements of children and their mothers collected in a randomized experiment studying the effect of home visits by specialists on future cognitive test scores of low birth weight, premature infants. The work [20] then "de-randomizes" the treatment assignment by removing children with non-white mothers from the treated group and simulates a treated and untreated outcome for each subject, allowing us to evaluate the performance of treatment effect estimation. We use the same data-preprocessing and evaluation strategy as in [14].

The large-scale database eICU has records of ICU patients from multiple hospitals across the continental United States [21]. We use the mean blood pressure as a corresponding outcome (Y) to the vasopressor as the treatment (T). Also, the patient features (X) include five static features (age, height, weight, race, and gender) and 25 dynamic features.

Among all hospitals we extract data of eight hospitals with the greatest data size to construct our dataset. Following [21], we use the sepsis cohort. We only select the data of each ICU admission within the first hour as time-dependency is not considered. The total number of selected records is 1,824. To synthesize the outcome given real treatment and patient features, we use the exact response surface B (except for the choice of ω_B) in IHDP [20] by which the outcome is in nonlinear response to covariates. Normalized patient features are used in the response surface below:

$$\begin{aligned}
y_i|T = 0 &\sim \mathcal{N}(\exp(x_i + W) \cdot \beta_B, 1) \\
y_i|T = 1 &\sim \mathcal{N}(x_i \cdot \beta_B - \omega_B, 1)
\end{aligned}$$

where $W = 0.5\ \mathbb{1}$, β_B is a vector of regression coefficients where each coefficient is randomly sampled from (0, 0.1, 0.2, 0.3, 0.4) with probabilities (0.6, 0.1, 0.1, 0.1, 0.1), which has the same dimension as X, and $\omega_B = 5$.

B. Baselines and Experimental Settings

We compare our model against the baselines below:

- **CT** [2] is a recently proposed causal tree model.
- **GANITE** [10] is a model based on generative adversarial networks (GANs) with counterfactual label generation during training.
- **CEVAE** [14] is one of the representative VAE-based models proposed to estimate causal effect.
- **TEDVAE** [15] is a VAE-based model with multiple disentangled latent variables incorporated.
- **R-Learner** [13] is a two-stage meta-learning algorithm with a nuisance component estimating propensity scores.
- **X-Learner** [11] is a two-step meta-learning algorithm which imputes treatment effect to augment the dataset.
- **CEMVAE-D** is a variant of the proposed model trained using the two-stage approach where the auxiliary models are disconnected from the backbone VAE model. In stage 1, Eq. 2 is used to train the auxiliary models until convergence, after which their parameters remain frozen. In stage 2, only Eq. 1 is used to train the backbone VAE model.

In our experiments, we use the same data replications from the GitHub page of [14] and follow the procedures according to [14] to evaluate the metrics (see Appendix C for more dataset details). To implement the causal tree and meta-learners, we use associated Python open packages (*CTL* for causal tree [2] and *causallib* for meta-learners [11], [13]). Since we use multi-layer perceptrons as the nuisance components for the meta-learners, we denote X-Leaner as X-MLP and R-Leaner as R-MLP. Empirical implementation shows that it takes less than five minutes for the most of time to train a CEMVAE model with one NVIDIA Ampere A100-80G GPU on eICU dataset.

During testing, to compute the outcomes of $p(Y|do(T), X)$, we average 250 samples from the approximate posterior $q(Z|X) = \int q(Z|X, H)q(H|X)dH$ where the variable H is a concatenation of $[Y_0, Y_1]$.

For performance evaluation, we adopt two commonly used metrics, namely Precision in Estimation of Heterogeneous Effect (ϵ_{PEHE}) and the absolute error of Average Treatment Effect (ϵ_{ATE}), defined as:

$$\epsilon_{\text{PEHE}} = \frac{1}{n} \sum_{i=1}^{n} (\tau(x_i) - \hat{\tau}(x_i))^2$$

TABLE I: Performance comparison on accuracy of treatment effect estimation based on three datasets.

Method	IHDP		eICU		Synthetic	
	$\sqrt{\epsilon_{\text{PEHE}}^{\text{within-s.}}}$	$\epsilon_{\text{ATE}}^{\text{within-s.}}$	$\sqrt{\epsilon_{\text{PEHE}}^{\text{within-s.}}}$	$\epsilon_{\text{ATE}}^{\text{within-s.}}$	$\sqrt{\epsilon_{\text{PEHE}}^{\text{within-s.}}}$	$\epsilon_{\text{ATE}}^{\text{within-s.}}$
CF	3.27±1.67	0.37±0.15	1.89±0.03	0.28±0.03	0.66±0.03	0.39±0.01
GANITE	6.86±2.59	4.25±0.69	7.46±0.48	1.25±1.10	0.95±0.03	0.25±0.17
CEVAE	2.20±1.07	0.22±0.10	1.88±0.07	0.03±0.01	0.34±0.01	0.09±0.01
X-MLP	1.83±0.49	**0.15±0.04**	1.74±0.10	0.47±0.16	0.36±0.03	0.07±0.05
R-MLP	2.93±1.19	0.40±0.09	1.87±0.30	0.18±0.09	0.40±0.04	0.08±0.05
TEDVAE	1.64±0.84	0.21±0.08	0.42±0.07	0.06±0.03	0.23±0.01	**0.02±0.01**
CEMVAE-D	1.56±0.74	0.27±0.08	0.62±0.08	0.30±0.10	0.29±0.06	0.09±0.09
CEMVAE	**1.39±0.63**	0.23±0.07	**0.41±0.02**	**0.01±0.01**	**0.21±0.01**	**0.02±0.01**
Method	$\sqrt{\epsilon_{\text{PEHE}}^{\text{out-of-s.}}}$	$\epsilon_{\text{ATE}}^{\text{out-of-s.}}$	$\sqrt{\epsilon_{\text{PEHE}}^{\text{out-of-s.}}}$	$\epsilon_{\text{ATE}}^{\text{out-of-s.}}$	$\sqrt{\epsilon_{\text{PEHE}}^{\text{out-of-s.}}}$	$\epsilon_{\text{ATE}}^{\text{out-of-s.}}$
CF	3.33±1.70	0.59±0.37	2.06±0.10	0.25±0.10	0.70±0.05	0.41±0.04
GANITE	6.54±2.25	4.56±.939	7.73±1.06	1.39±1.28	1.16±0.02	0.26±0.17
CEVAE	2.75±1.70	0.25±0.07	1.78±0.14	0.08±0.04	0.33±0.02	0.09±0.02
X-MLP	1.63±0.30	0.23±0.04	1.82±0.07	0.40±0.10	0.38±0.04	0.08±0.05
R-MLP	2.62±0.92	0.39±0.11	1.63±0.28	0.14±0.14	0.38±0.02	0.07±0.05
TEDVAE	1.33±0.58	0.23±0.09	0.42±0.02	0.07±0.04	0.23±0.01	0.02±0.01
CEMVAE-D	1.39±0.52	0.25±0.07	0.57±0.09	0.31±0.10	0.29±0.04	0.10±0.09
CEMVAE	**1.23±0.42**	**0.16±0.06**	**0.36±0.08**	**0.03±0.02**	**0.22±0.01**	**0.01±0.00**

$$\epsilon_{\text{ATE}} = \frac{1}{n}\sum_{i=1}^{n}\tau(x_i) - \frac{1}{n}\sum_{i=1}^{n}\hat{\tau}(x_i)$$

where $\tau(x_i)$ refers to the ground truth treatment effect of a subject, $\hat{\tau}(x_i)$ is the estimated treatment effect, and n is the size of test set for a target subgroup.

For all the baselines and our proposed model, we carefully conduct the grid search on their hyper-parameters and select the best results among the combinations of hyper-parameters based on the evaluation metrics.

C. Results and Analysis

We summarize the performance of the baselines and the proposed method for treatment effect estimation in Table I for comparison. The empirical results show that CEMVAE obtains the best performance in terms of treatment effect estimation as measured by $\sqrt{\epsilon_{\text{PEHE}}}$ and ϵ_{ATE} for most datasets.

To show the merits of the data augmentation approach in learning a more latent variable which can generalize better, we plot the distributions of Z given the training and test data, as shown in Fig. 3. According to the figure, we observe that CEVAE has the greatest Wasserstein distance of Z between training and test data, which can be attributed to its switching mechanism that the variables of subjects from control and treatment groups are separately encoded. This results in that only the estimation head $q(y|t,z)$ is trained during training. In contrast, both $q(y|t,z)$ and $q(y|1-t,z)$ of CEMVAE are trained during training. TEDVAE can give a smaller Wasserstein distance as the variables of subjects from both control and treatment groups are encoded by the same encoders and its disentangled factors of Z reduce the selection bias. CEMVAE gives the least Wasserstein distance because it learns unbiased laten representations. Since $P(X|T=0) = P(X|T=1)$ given $\mathcal{D}_{obs} \bigcup \mathcal{D}_{cf}$, a sufficient condition for $q(Z|X, Y_t, Y_{1-t})$ to be the same between $t=0$ and $t=1$ is that the counterfactual outcome Y_{1-t} becomes oracle estimation, which indicates that the selection bias can be reduced between groups of treatment and control. In other words, for CEMVAE, the challenges of both generating accurate Y_{1-t} and reducing selection bias via learning unbiased Z can be addressed simultaneously by only addressing the former. In contrast, maximizing the component of $\log q(T|Z)$ in the learning objectives by both CEVAE and TEDVAE can be counterproductive to reduce selection bias as the discrepancy of latent representations between groups of treatment and control is preserved. Additionally, the size of the treatment group and that of control group are more balanced for CEMVAE, while both CEVAE and TEDVAE are trained with only observational data in which the size of the treatment group is much larger than that of the control group.

To show that the accuracy of the imputed values can be improved by the EM algorithm, we compare CEMVAE with CEMVAE-D which is trained in a two-stage fashion. To evaluate the accuracy for testing, we calculate the *out-of-sample* RMSE:

$$\textit{RMSE-OOS} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\sum_{t=0}^{1}[y_i - \mathbb{E}[(\hat{y}_i])^2|x_i, \phi_t, t]}$$

where n is the number of data points, y_t is the ground truth outcome, and $\hat{y}_t$ is the expected value calculated by $q_{\phi_t}(y_{t_i,i}|x_i, T = t_i)$. For the accuracy of the generated counterfactuals during training, we calculate *within-sample* RMSE:

$$\textit{RMSE-Within} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[(y_i - \mathbb{E}[(\hat{y}_i])^2|x_i, \phi_{1-t_i}, 1-t_i]}$$

where t_i is a factual treatment in the training set.

According to Table II, it is observed that CEMVAE can give a significantly lower RMSE than CEMVAE-D by a large margin for all three datasets. This shows that the end-to-end learning framework can improve the accuracy of imputed counterfactuals on both training and test sets.

TABLE II: Comparison between end-to-end learning and the setting with disconnected counterfactual generators.

Method	IHDP		eICU		Synthetic	
	RMSE-OOS	*RMSE-Within*	*RMSE-OOS*	*RMSE-Within*	*RMSE-OOS*	*RMSE-Within*
GANITE	—	5.89±2.16	—	8.19±8.49	—	0.95±0.06
X-MLP	1.39±0.27	1.55±0.32	1.58±0.09	1.44±0.03	0.30±0.01	0.35±0.02
CEMVAE-D	1.05±0.34	1.12±0.43	0.20±0.02	0.20±0.00	0.20±0.01	0.24±0.00
CEMVAE	**0.94±0.42**	**0.68±0.25**	**0.18±0.02**	**0.17±0.01**	**0.18±0.01**	**0.22±0.00**

VI. Conclusion

This paper proposes an end-to-end framework for jointly learning a pseudo counterfactual generator and a VAE-based model to reduce the bias in treatment effect estimation. We formulate it as a missing data problem and reparameterize the VAE model to be trained using the EM algorithm. The proposed model named CEMVAE can generate counterfactuals of higher quality leading to improved accuracy on estimating the treatment effect. Rigorous empirical evaluation has been carried out based on synthetic and semi-synthetic datasets to demonstrate its superior performance over the SOTA methods. Furthermore, CEMVAE is proposed for binary treatment with the main objective to evaluate the effectiveness of the end-to-end learning framework. Extending the framework to continuous treatment effect estimation with dosage level considered will be an interesting future direction.

Acknowledgements

This research is partially supported by the Research Matching Grant Scheme RMGS2021_8_06 from the Hong Kong Government.

References

[1] Patrick Schwab, Lorenz Linhardt, Stefan Bauer, Joachim M. Buhmann, and Walter Karlen. Learning counterfactual representations for estimating individual dose-response curves. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 5612–5619, April 2020.

[2] Christopher Tran and Elena Zheleva. Learning triggers for heterogeneous treatment effects. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 5183–5190, July 2019.

[3] Fredrik Johansson, Uri Shalit, and David Sontag. Learning representations for counterfactual inference. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 3020–3029, New York, New York, USA, 20–22 Jun 2016. PMLR.

[4] Alexander Peysakhovich and Akos Lada. Combining observational and experimental data to find heterogeneous treatment effects. *ArXiv*, abs/1611.02385, 2016.

[5] Yuhao Wang and Rajen Dinesh Shah. Debiased inverse propensity score weighting for estimation of average treatment effects with high-dimensional confounders. *arXiv: Methodology*, 2020.

[6] Paul R. Rosenbaum and Donald B. Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 04 1983.

[7] Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 01 2018.

[8] Ioana Bica, Ahmed M Alaa, James Jordon, and Mihaela van der Schaar. Estimating counterfactual treatment outcomes over time through adversarially balanced representations. In *Proceedings of International Conference on Learning Representations*, 2020.

[9] Valentyn Melnychuk, Dennis Frauen, and Stefan Feuerriegel. Causal transformer for estimating counterfactual outcomes. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 15293–15329. PMLR, 17–23 Jul 2022.

[10] Jinsung Yoon, James Jordon, and Mihaela van der Schaar. GANITE: Estimation of individualized treatment effects using generative adversarial nets. In *Proceedings of International Conference on Learning Representations*, 2018.

[11] Sören R. Künzel, Jasjeet S. Sekhon, Peter J. Bickel, and Bin Yu. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences*, 116(10):4156–4165, 2019.

[12] Dennis Frauen and Stefan Feuerriegel. Estimating individual treatment effects under unobserved confounding using binary instruments. In *The Eleventh International Conference on Learning Representations*, 2023.

[13] Xinkun Nie and Stefan Wager. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319, 09 2020.

[14] Christos Louizos, Uri Shalit, Joris Mooij, David Sontag, Richard Zemel, and Max Welling. Causal effect inference with deep latent-variable models. In *Advances in Neural Information Processing Systems*, page 6449–6459, Red Hook, NY, USA, 2017. Curran Associates Inc.

[15] Weijia Zhang, Lin Liu, and Jiuyong Li. Treatment effect estimation with disentangled latent factors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 10923–10930, May 2021.

[16] PAUL R. ROSENBAUM and DONALD B. RUBIN. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 04 1983.

[17] Susan Athey and Guido Imbens. Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27):7353–7360, 2016.

[18] Yu Gong, Hossein Hajimirsadeghi, Jiawei He, Thibaut Durand, and Greg Mori. Variational selective autoencoder: Learning from partially-observed heterogeneous data. In Arindam Banerjee and Kenji Fukumizu, editors, *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, volume 130 of *Proceedings of Machine Learning Research*, pages 2377–2385. PMLR, 13–15 Apr 2021.

[19] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *International Conference on Learning Representations (ICLR)*, 2014.

[20] Craig T. Ramey, Donna M. Bryant, Barbara Hanna Wasik, Joseph Sparling, Kaye Fendt, and Lisa M LaVange. Infant health and development program for low birth weight, premature infants: program elements, family participation, and child intelligence. *Pediatrics*, 89 3:454–65, 1992.

[21] Tom Pollard, Alistair Johnson, Jesse Raffa, Leo Celi, Roger Mark, and Omar Badawi. The eicu collaborative research database, a freely available multi-center database for critical care research. *Scientific Data*, 5:180178, 09 2018.