Towards Effective and Efficient Multi-valued Treatment Uplift Modeling in Online Marketing

Zexu Sun[1,2], Dugang Liu[3(✉)], Xing Tang[4], Yunpeng Weng[4], and Xiuqiang He[4]

[1] Gaoling School of Aritficial Intelligence, Renmin University of China, Beijing, China
sunzexu21@ruc.edu.cn
[2] Beijing Key Laboratory of Big Data Management and Analysis Methods, Beijing, China
[3] College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China
dugang.ldg@gmail.com
[4] FiT, Tencent, Shenzhen, China
{shawntang,edwinweng,xiuqianghe}@tencent.com

Abstract. The effectiveness of an online marketing campaign heavily relies on the identification of user groups that exhibit high sensitivity to specific treatments. However, existing works in this domain has encountered certain limitations when applied in practical settings. Firstly, most studies have primarily focused on binary treatment scenarios, but real-world industrial applications often involve multi-valued treatments, rendering these approaches incompatible. Secondly, although a few studies have addressed multi-valued treatment scenarios, many of them have directly extended binary treatment architectures without considering additional optimization. This oversight can result in redundant model parameters and performance bottlenecks. In order to encounter aforementioned challenges, we propose a novel reparameterization multi-head treatment uplift network, or RMNet for short. RMNet incorporates an invariant feature representation and a reparameterization multi-head module. This module achieves a balanced representation of all treatments by employing gradient constraints, thereby mitigating selection bias and enhancing model efficiency and performance. The latter responses to different treatments as offsets relative to the control response, thus we employ a reparameterization multi-head structure to effectively reduce the number of model parameters necessary for predicting responses to different treatments. Finally, extensive experiments are conducted on two datasets to demonstrate the effectiveness and efficiency of our RMNet.

Keywords: Uplift modeling · Online Marketing · Multi-valued Treatment · Reparameterization

M. Onizuka et al. (Eds.): DASFAA 2024, LNCS 14856, pp. 121–138, 2024.
https://doi.org/10.1007/978-981-97-5575-2_8

1 Introduction

Enhancing user engagement and increasing platform revenue through effective online marketing strategies is a critical objective of the online platform [8]. These strategies encompass the implementation of well-designed incentives, including coupons, discounts, and bonuses. For the successful implementation of an online marketing strategy that ensures efficient delivery and minimal costs, precise identification of the target user group for each incentive is more important. To achieve this goal, a marketing model needs to discern the impact of different incentives on user response and focus on high-gain users within each incentive. In practical scenarios, we often observe only one type of user response, which can be attributed to either a specific incentive (*i.e.*, treatment group) or without any incentive (*i.e.*, control group). This problem is different from traditional supervised learning tasks and is commonly regarded as a causal inference problem, where the estimation of the impact of different incentives on user responses is regarded as an individual treatment effect (ITE) [16], also known as uplift. As a result, previous studies have proposed various uplift modeling techniques and demonstrated their effectiveness in online marketing [6,12].

The existing body of uplift modeling methods primarily centers around three research directions: 1) Meta-learner-based approaches: This research direction leverages the off-the-shelf estimators as base learners for different groups. The predicted differences between these groups are then interpreted as uplifts. Notably, the S-Learner [7] and T-Learner [7] are two representative methods within this direction. They employ a single global estimator or estimators corresponding to different treatment types. 2) Tree-based (or Forest-based) approaches: The fundamental concept behind this research direction involves partitioning users into distinct subgroups based on their uplift values within the feature space. Uplift prediction can be performed on each leaf node by employing diverse splitting criterias [12,15]. Notably, the Causal Forest [18] employs an ensemble of multiple trees to estimate treatment effect. 3) Neural network-based approaches: This research direction leverages neural networks with flexible structures to predict uplift. Consequently, it is natural to adapt neural networks as base learners for meta-learner methods. In recent years, some representation learning based methods are proposed, such as TARNet [13], CFRNet [13], Dragonnet [14], and others, which aim to mitigate various biases prevalent in online marketing scenarios. In this study, we specifically focus on the neural network-based research direction due to its enhanced flexibility and generalizability compared to other research directions.

Although the existing uplift models for online marketing have shown promising results, most have focused on the binary treatment setting, which only considers whether to deliver a coupon to a user. However, in many real-world scenarios, a marketing campaign is usually more likely to involve multiple incentives. For example, a service platform's discount coupons may be of multiple different amounts. To solve this problem, some recent works have focused on addressing the setting of multi-valued treatments, *e.g.*, MEMENTO [9] and HydraNet [17]. However, these methods are usually directly extended based on the represen-

tative architecture of the above-mentioned binary treatment and do not make more additional optimizations for the characteristics of multi-valued treatment. This will make these methods likely to have redundant model parameters and bottlenecks in performance and efficiency. For ease of understanding, we show in Fig. 1 the architecture of representative multi-valued treatment methods and some binary methods extended to the multi-valued treatment setting. We can see that existing methods design the same network structure for each response prediction, and this will significantly lead to excessive parameter size and inefficient training when the number of treatments increases. In addition, existing methods usually adopt alignment constraints on different treatment representations (*e.g.*, various distribution distances) to solve the selection bias, and this approach is also prone to high complexity when the number of treatments increases [20].

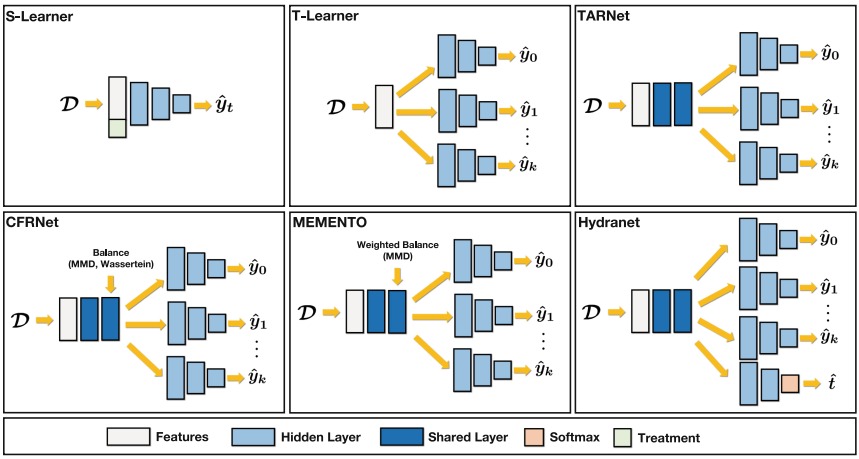


Fig. 1. Architecture diagram of some representative neural network-based uplift modeling methods with multi-valued treatment, where $\mathcal{D}$ represents the training set with $\{(\boldsymbol{x}_i, t_i, y_i)\}_{i=1}^{n}$, $\hat{y}_0, \hat{y}_1, \ldots, \hat{y}_k$ denote the predicted responses for the control group and treatment groups, respectively, and $\hat{t}$ is the predicted treatment.

To address the aforementioned challenges, we propose a novel reparameterization multi-head treatment uplift network, named RMNet. RMNet consists of two customized modules designed to alleviate the limitations of existing methods. Firstly, an invariant feature representation module is employed to obtain a balanced representation for all treatments. This is achieved by treating each treatment as a domain and utilizing domain-invariant learning with gradient constraints. Through this, we effectively reduce the complexity associated with aligning different treatment representations, especially when the number of treatment increases. Secondly, a reparameterization multi-head module is introduced, which regards the responses to different treatments as offsets based on the control response (*i.e.*, without any treatment). A reparameterization multi-head structure is utilized to predict the response to different treatments. This approach

significantly reduces the number of model parameters required for prediction heads for different treatments and enhances knowledge transfer between treatments. Furthermore, we provide a theoretical analysis of the invariant balanced representation, highlighting its significance in decision-making. Additionally, we conduct comprehensive offline and online evaluations, the experimental results demonstrate the effectiveness and efficiency of our proposed RMNet.

2 Related Works

Uplift Modeling with Binary Treatment. In binary treatment scenario, we review the related works by the three research lines mentioned in the above. 1) S-Learner [7] and T-Learner [7] directly model the user response by using one or two base learners, respectively. To mitigate the selection bias, X-Learner [7] and R-Learner [10] are proposed to get unbiased and more accurate uplift prediction. 2) Uplift-Tree [12] is modeled on the standard decision tree to continuously select the optimal split features and split points according to the size of the information gain, then to achieve the process of precise stratification. The core idea is to design the splitting criterion to make the prediction of uplift after splitting more accurate. Especially, Causal Forest [18] ensemble multiple causal trees to predict the uplift. Based on this structure, by designing different split criterions, there are some variants [20] which can get more accurate uplift prediction in various scenarios. 3) Neural network based methods have got great research interest in recent years. TARNet [13] constructs a shared feature representation and two seperate response prediction heads for the uplift prediction. On the basis of it, CFRNet [13] introduces the integral probability metrics (IPM) to train a balanced representation for treatment and control groups. Dragonnet [14] leverages the target regularization approach to model the prediction error of the separate label prediction heads. Most of the uplift models designed for binary treatment are difficult to extend to the multi-valued treatment setting. Our RMNet is a novel uplift model designed for multi-valued treatment.

Uplift Modeling with Multi-valued Treatment. Due to the uplift modeling problems with multi-valued treatment are common in real-world applications, some works [4] are proposed to solve this problem with different model structures. [21] extends standard meta learner-based uplift models to support multi-valued treatment with different costs. On top of CFRNet, [9] propose MEMENTO to get a balanced representation for different treatment and control groups. [17]) propose Hydranet, which reform the dragonnet to support for the multi-valued treatment setting. For the extension of meta learners, [1] carry out a theoretical analysis of their error upper bounds as functions of important parameters, and show that the naive extensions do not always provide satisfactory results. However, the proposed neural network-based methods for multiple treatment uplift modeling always have numerous model parameters, which decrease the model efficiency. Different from above works, our RMNet achieves more accurate uplift predictions using fewer model parameters, resulting in improved model efficiency.

3 Preliminaries

Let $\mathcal{D} = \{(\boldsymbol{x}_i, t_i, y_i)\}_{i=1}^n$ represents the observed dataset consisting of n samples. Without loss of generality, $\boldsymbol{x}_i \in \mathcal{X} \subset \mathbb{R}^d$ denotes a d-dimensional feature vector. $y_i \in \mathcal{Y}$ represents the response variable, where $\mathcal{Y}$ can be either binary or continuous. $t_i \in \mathcal{T} \in \{0, 1, \ldots, K\}$ with $K \geq 2$ is the treatment indicator variable, indicating different treatments such as various discount levels. Note that $t_i = 0$ corresponds to the situation where no treatment is applied. To formalize the problem, we adopt the Neyman-Rubin potential outcome framework [11] to define the uplift modeling problem with multi-valued treatment. Let $y_i(k)$ and $y_i(0)$ denote the potential outcomes for user i when receiving treatment $t_i = k \in \{1, \ldots, K\}$ or not receiving any treatment, respectively. The incremental improvement caused by a specific treatment k (referred to as individual treatment effect or uplift) is denoted as τ_i^k, defined as:

$$\tau_i^k = y_i(k) - y_i(0). \tag{1}$$

Since we can only observe one response (*i.e.*, $y_i(k)$ or $y_i(0)$) for each user, we do not have access to the actual uplift label τ_i^k. Fortunately, under certain appropriate assumptions [16], we can use the conditional average treatment effect (CATE) as an unbiased estimator for the uplift. CATE is defined as:

$$\begin{aligned} \tau_i^k(\boldsymbol{x}_i) &= \mathbb{E}(y_i(k) - y_i(0)|\boldsymbol{x}_i) \\ &= \underbrace{\mathbb{E}(y_i(k)|t_i = k, \boldsymbol{x}_i)}_{\mu_k(\boldsymbol{x}_i)} - \underbrace{\mathbb{E}(y_i(0)|t_i = 0, \boldsymbol{x}_i)}_{\mu_0(\boldsymbol{x}_i)}. \end{aligned} \tag{2}$$

This represents the expected treatment effect between $t_i = 0$ and $t_i = k$ given $\boldsymbol{x}_i$. For brevity of notation, we will omit the subscript i in the following if there is no ambiguity. Intuitively, the desired objective can be expressed as the difference between two conditional means $\tau^k(\boldsymbol{x}) = \mu_k(\boldsymbol{x}) - \mu_0(\boldsymbol{x})$. Once we obtain the uplift predictions $\hat{\tau}^k(\boldsymbol{x})$ for all k different incentives, we can use an evaluation metric to rank users and make the final decision regarding treatment assignment.

4 Methodology

In the high level, we aim to design a more effective and efficient uplift model for multi-valued treatment setting, which incorporates an invariant feature representation module and a reparameterization multi-head module.

4.1 Architecture

The proposed model, called the Reparameterization Multi-Head Treatment Uplift Network with Invariant Representation (RMNet), is depicted in Fig. 2. Given a sample $\{\boldsymbol{x}, t, y\}$, the invariant feature representation module employs established feature encoding techniques to convert the feature vectors $\boldsymbol{x}$ into

embedding representations. We refer to this process as $\Phi(\cdot)$ and denote the resulting embedding representation as $\Phi(\boldsymbol{x})$. In contrast to most existing methods that address selection bias by imposing an alignment constraint on the embedded representations across different treatment groups, we treat different treatments as distinct domains. This approach transforms the problem into a domain-invariant representation, *i.e.* $\Phi(\boldsymbol{x}) \perp\!\!\!\perp t$. Then, following the acquisition of the balanced representation from the previous module, our reparameterization multi-head module treats the responses of different treatments as offsets based on the control response. It introduces a reparameterization structure for prediction. Instead of separately estimating $\mu_0(\boldsymbol{x})$ and $\mu_k(\boldsymbol{x})$, we build upon the identity $\mu_k(\boldsymbol{x}) = \mu_0(\boldsymbol{x}) + \tau^k(\boldsymbol{x}), k \in \{1, \ldots, K\}$. The offset $\mu_0(\boldsymbol{x})$ directly contributes to the control response prediction head h_0, while τ^k is estimated by the uplift prediction heads h_{τ^k}. Moreover, h_0 can be designed with a more complex structure compared to h_{τ^k} to capture shared information among different treatment response predictions. To further discourage redundancy and encourage the identification of private structures, we apply a regularizer to orthogonalize the shared control response space and the private subspaces of different treatments. For training our RMNet, all prediction heads are jointly optimized with parameter complexity regularization. The final learning objective of our RMNet can be formulated as follows:

$$\min_{\Theta_f, \Theta_t, \Theta_{h_0}, \Theta_{h_{\tau^k}}} \mathcal{L}_{RMNet} = \mathcal{L}_y + \alpha(\mathcal{L}_t^{\Theta_t} - \mathcal{L}_t^{\Theta_f}) + \lambda_1 R_o + \lambda_2 (R_\Phi + R_{h_0} + \textstyle\sum_{k \in \{1,\ldots,K\}} R_{h_{\tau^k}}), \tag{3}$$

where $\mathcal{L}_y = L\left(y, h_0 + h_{\tau^k}\right)$ denotes the response prediction loss, and for continuous y, $L(\cdot,\cdot)$ is the Mean Square Error (MSE) loss, otherwise $L(\cdot,\cdot)$ is the Binary Cross Entropy (BCE) loss. The prediction loss $\mathcal{L}_t$ for the treatment prediction head requires performing a gradient inversion operation. R_Φ, R_{h_0}, and $R_{h_{\tau^k}}$ represent the parameter regularization terms, typically implemented as L2 regularization, for the balanced representation, shared reparameterization head, and separate uplift prediction heads, respectively. R_o denotes the orthogonal regularization for the control prediction head h_0 in relation to the other uplift prediction heads h_{τ^k}. The hyper-parameters α, λ_1, and λ_2 control the trade-off between different components. Next, we provide a detailed description of each module based on the training process.

4.2 Invariant Feature Representation Module

To begin, we consider K different potential treatments, denoted as $k \in \{1, \ldots, K\}$, and one control group, which represent distinct domains. Our objective is to construct a balanced representation of features $\boldsymbol{x}$ that exhibits an invariant distribution across different domains:

$$P\left(\Phi\left(\boldsymbol{x}\right) \mid t = 0\right) = \cdots = P\left(\Phi\left(\boldsymbol{x}\right) \mid t = k\right). \tag{4}$$

To facilitate explanation, we denote the treatment prediction as $Q_t(\Phi(\boldsymbol{x}); \Theta_t)$ and the response prediction as $Q_y(\Phi(\boldsymbol{x}); \Theta_y)$. Here, $\Theta_y = \{\Theta_{h_0}, \Theta_{h_{\tau^1}}, \ldots, \Theta_{h_{\tau^k}}\}$

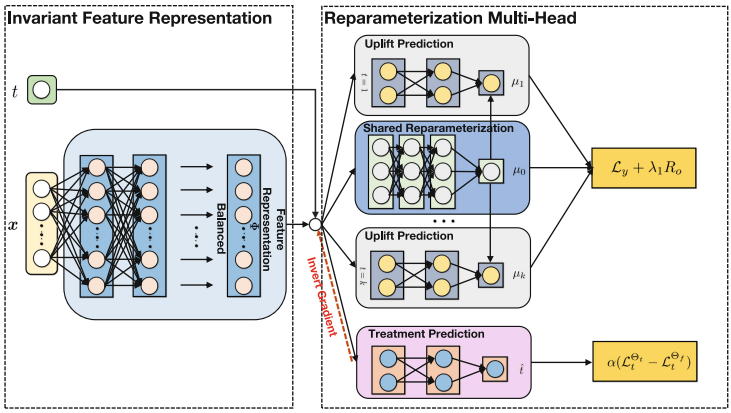


Fig. 2. The overall architecture of our RMNet. In addition, we also add the L2 regularization for each neural network in the architecture, which is shown in Eq. (3).

represents the parameters for the response prediction heads. Additionally, the feature representation network $\Phi(\cdot)$ is parameterized by Θ_f, denoted as $\Phi(\boldsymbol{x}; \Theta_f)$. Consequently, for each sample in the dataset, the loss for response prediction can be defined as:

$$\mathcal{L}_y = \| y - Q_y \left(\Phi \left(\boldsymbol{x}; \Theta_f \right); \Theta_y \right) \|^2. \tag{5}$$

Note that the BCE loss can be used for the binary response instead of the MSE loss. Similarly, for the treatment prediction, the loss function $\mathcal{L}_t$ is given by:

$$\mathcal{L}_t = -\sum_{j=0}^{K} \mathbb{I}_{t=j} \log \left(Q_t \left(\Phi \left(\boldsymbol{x}; \Theta_f \right); \Theta_t \right) \right). \tag{6}$$

Subsequently, in order to construct a treatment-aware balanced representation from the standpoint of domain-invariant representation, our objective is to minimize the encoding of domain information in feature representations, while simultaneously minimizing the losses in response prediction. In other words, the feature representations should not enable accurate restoration of domain labels. Thus, the overall loss $\mathcal{L}_b$ for balanced representation can be summarized as:

$$\mathcal{L}_b(\Theta_f, \Theta_t, \Theta_y) = \mathcal{L}_y(\Theta_f, \Theta_y) - \alpha \mathcal{L}_t(\Theta_f, \Theta_t), \tag{7}$$

where α is the hyper-parameter to control the trade-off between treatment and response predictions. Through the minimization of the aforementioned objective, we can acquire a balanced representation that effectively mitigates the selection bias present in the dataset. However, in practical implementation, directly solving Eq. (7) poses an uncontrollable risk. This is due to the fact that the optimization of the second term on the right-hand side may lead the model to opportunistically optimize its performance, subsequently impacting the efficacy of each subsequent treatment prediction head. To address this challenge and efficiently solve Eq. (7), we can reformulate the objective function $\mathcal{L}_b(\Theta_f, \Theta_t, \Theta_y)$

as the search for a saddle point $\left(\hat{\Theta}_f, \hat{\Theta}_t, \hat{\Theta}_y\right)$ that achieves equilibrium between treatment and response prediction.

$$\begin{aligned} \left(\hat{\Theta}_y, \hat{\Theta}_f\right) &= \arg \min_{\hat{\Theta}_f, \Theta_y} \mathcal{L}_y\left(\Theta_f, \Theta_y\right) - \alpha \mathcal{L}_t(\Theta_f, \hat{\Theta}_t), \\ \hat{\Theta}_t &= \arg \max_{\Theta_t} \mathcal{L}_y\left(\hat{\Theta}_f, \hat{\Theta}_y\right) - \alpha \mathcal{L}_t(\hat{\Theta}_f, \Theta_t). \end{aligned} \tag{8}$$

From an intuitive perspective, the aforementioned operations can often be incorporated within an uplift network structure, situated between two components: response predictions that anticipate the target labels, and treatment predictions that forecast the treatments received by the user. Drawing inspiration from prior research on domain-invariant representations [5], we can effectively optimize the aforementioned objective by reversing the gradient direction on the feature representation. This reversal occurs during the gradient backpropagation process of the model and is driven by the loss associated with treatment prediction. Consequently, the feature representation receives gradients aimed at minimizing the loss in response prediction, while simultaneously maximizing the loss in treatment prediction. This approach ensures a balanced representation. To achieve the desired stationary point as outlined in Eq. (8), we employ the following gradient update step.

$$\begin{aligned} \Theta_f &\longleftarrow \Theta_f - \left(\frac{\partial \mathcal{L}_y}{\partial \Theta_f} - \alpha \frac{\partial \mathcal{L}_t}{\partial \Theta_f}\right), \\ \Theta_y &\longleftarrow \Theta_y - \frac{\partial \mathcal{L}_y}{\partial \Theta_y} \\ \Theta_t &\longleftarrow \Theta_t - \alpha \frac{\partial \mathcal{L}_t}{\partial \Theta_t} \end{aligned} \tag{9}$$

In our experimental setup, we initially assign an arbitrary value to α, and during the training process, we apply an exponentially increasing schedule.

Finally, we provide corresponding theoretical insights into the above process to show that it can obtain the desired balanced representation, *i.e.*, Theorem 1.

Theorem 1. *For each $k \in \{0, 1, \ldots, K\}$, we denote the distribution of $\Phi(\boldsymbol{x}; \Theta_f)$ conditioned on $t = k$ as P_k^{Φ}. Additionally, let Q_t^k represent the output of Q_t corresponding to treatment k. The minimax game can be formally defined as follows:*

$$\begin{aligned} &\min_{\Theta_f} \max_{\Theta_t} \sum_{k=0}^{K} \mathbb{E}_{\boldsymbol{x}}\left[\log\left(Q_t^k\left(\Phi\left(\boldsymbol{x}; \Theta_f\right); \Theta_t\right)\right)\right], \\ &\quad s.t. \sum_{k=0}^{K} Q_t^k\left(\Phi\left(\boldsymbol{x}; \Theta_f\right)\right) = 1. \end{aligned} \tag{10}$$

The global minimum of the function is achieved if and only if the learned representations exhibit invariance across all domains. This can be denoted as $P_0^{\Phi} = P_1^{\Phi} = \ldots = P_K^{\Phi}$.

To prove this theorem, we first prove the following proposition.

Proposition 1. *Given a fixed* Φ*, we define* $\boldsymbol{x}' = \Phi(\boldsymbol{x})$*. In this case, the optimal prediction probabilities of* Q_t *can be expressed as follows:*

$$Q_t^{j^*}(\boldsymbol{x}') = \frac{P_j^{\Phi}(\boldsymbol{x}')}{\sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}')}.$$

Proof. For fixed Φ, the optimal prediction probabilities are given by

$$Q_t^* = \arg\max_{\Theta_t} \sum_{j=1}^{K} \int_{\boldsymbol{x}'} \log\left(Q_t^j(\boldsymbol{x}')\right) P_j^{\Phi}(\boldsymbol{x}')\, d\boldsymbol{x}' \qquad \text{s.t.} \sum_{j=1}^{K} Q_t^j(\boldsymbol{x}') = 1.$$

Maximizing the value function pointwise and applying Lagrange multiplies, we get

$$Q_t^* = \arg\max_{\Theta_t} \sum_{j=1}^{K} \log\left(Q_t^j(\boldsymbol{x}')\right) P_j^{\Phi}(\boldsymbol{x}') + \alpha\left(\sum_{j=1}^{K} Q_t^j(\boldsymbol{x}') - 1\right).$$

By setting the derivative with respect to $Q_t^{j^*}(\boldsymbol{x}')$ to zero and solving for $Q_t^{j^*}(\boldsymbol{x}')$, we obtain the following expression:

$$Q_t^{j^*}(\boldsymbol{x}') = -\frac{P_j^{\Phi}(\boldsymbol{x}')}{\alpha}$$

Here, the value of α can be determined by satisfying the constraint $\alpha = -\sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}')$.

Next, we can get the proof of Theorem 1.

Proof. By substituting the expression derived from Proposition 1 into the minimax game defined in Eq. (10), the objective function for Φ can be formulated as follows:

$$\min_{\Phi} \sum_{j=0}^{K} \mathbb{E}_{\boldsymbol{x}' \sim P_j^{\Phi}}\left[\log\left(\frac{P_j^{\Phi}(\boldsymbol{x}')}{\sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}')}\right)\right]$$

We then note that

$$\begin{aligned}
&\sum_{j=0}^{K} \mathbb{E}_{\boldsymbol{x}' \sim P_j^{\Phi}} \left[\log \left(\frac{P_j^{\Phi}(\boldsymbol{x}')}{\sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}')} \right) \right] + K \log K \\
&= \sum_{j=0}^{K} \left(\mathbb{E}_{\boldsymbol{x}' \sim P_j^{\Phi}} \left[\log \left(\frac{P_j^{\Phi}(x')}{\sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}')} \right) \right] + \log K \right) \\
&= \sum_{j=0}^{K} \mathbb{E}_{\boldsymbol{x}' \sim P_j^{\Phi}} \left[\log \left(\frac{P_j^{\Phi}(\boldsymbol{x}')}{\frac{1}{K} \sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}')} \right) \right] \\
&= \sum_{j=0}^{K} KL \left(P_j^{\Phi}(\boldsymbol{x}') \| \frac{1}{K} \sum_{i=0}^{K} P_i^{\Phi}(\boldsymbol{x}') \right) \\
&= K \cdot JSD\left(P_0^{\Phi}, \ldots, P_K^{\Phi} \right)
\end{aligned}$$

where $KL(\cdot\|\cdot)$ denotes the Kullback-Leibler divergence, while $JSD(\cdot, \ldots, \cdot)$ represents the multi-distribution Jensen-Shannon Divergence. It should be emphasized that $K \log K$ is a constant, and the multi-distribution JSD is always non-negative, attaining a value of 0 only when all the distributions are identical. Consequently, we can deduce that $P_0^{\Phi} = P_1^{\Phi} = \ldots = P_K^{\Phi}$.

4.3 Reparameterization Multi-head Module

Instead of estimating $\mu_0(\boldsymbol{x})$ and $\mu_k(\boldsymbol{x})$ separately, we build upon the identity $\mu_k(\boldsymbol{x}) = \mu_0(\boldsymbol{x}) + \tau^k(\boldsymbol{x})$, where $k \in \{1, \ldots, K\}$. We utilize a shared reparameterization head responsible for predicting control responses (i.e., responses without any treatment), while other treatment prediction heads are responsible for predicting uplifts associated with different treatments. This modified network structure effectively reduces the number of model parameters compared to the existing structure. The shared reparameterization head facilitates better knowledge transfer among different prediction heads, thereby mitigating selection bias in the data. Furthermore, as illustrated in Fig. 2, we can design a more complex structure for the shared reparameterization head compared to the lightweight prediction heads. This enhanced structure enables better capturing of shared information within the dataset. Specifically, we denote the shared reparameterization head as h_0 and each uplift prediction head as h_{τ^k}. Consequently, the loss function $\mathcal{L}_y$ in Eq. (3) can be formulated as follows:

$$\mathcal{L}_y = \|y - h_0(\Phi(\boldsymbol{x}))\|^2 + \sum_{k \in \{1, \cdots, K\}} \|y - (h_0(\Phi(\boldsymbol{x})) + h_{\tau^k}(\Phi(\boldsymbol{x})))\|^2. \tag{11}$$

Note that the BCE loss can be used for the binary response instead of the MSE loss.

Moreover, we incorporate an extra orthogonality penalty between the shared reparameterization head and the other treatment prediction heads. This penalty

serves to discourage redundancy and promote the identification of distinct structures.

In particular, we employ an orthogonality constraint (R_o) between the subspaces of each uplift prediction head and the shared parameterization prediction head at every layer.

$$R_o\left(\Theta_{h_0}, \Theta_{h_{\tau^k}}\right) = \sum_{k\in\{1,\cdots,K\}} \sum_{l=1}^{L} \left\| \theta_{h_0}^{l\top} \theta_{h_{\tau^k},1:m_s^{l-1}} \right\|_F^2, \tag{12}$$

where $\|\cdot\|_F^2$ represents the squared Frobenius norm. The variable l denotes the layer number, where $l > 1$. The terms p_h^{l-1} and p_τ^{l-1} refer to the output dimensions of the shared and private subspaces in the previous layer, respectively. The weights in the shared subspace are denoted as $\theta_h^l \in \mathbb{R}^{p_h^{l-1}\times p_\tau^l}$, while $\theta_{h_{\tau^k}}^l \in \mathbb{R}^{\left(p_h^{l-1}+p_\tau^{l-1}\right)\times p_h^l}$ represents the weights in each private subspace. The hyper-parameter λ_2 is used to control the trade-off. To promote orthogonality or lack of correlation between the weight matrices of different prediction heads in our RMNet, we incorporate this regularization term into the final loss function during training. Furthermore, in order to regularize the complexity of $\Phi(\cdot)$, h_0, and each h_{τ^k} individually, we apply L2 regularization $R(\cdot)$ to the parameters of all the prediction heads.

$$R_\Phi + R_{h_0} + \sum_{k\in\{1,\ldots,k\}} R_{h_{\tau^k}} = R(\Theta_f) + R(\Theta_{h_0}) + \sum_{k\in\{1,\ldots,K\}} R(\Theta_{h_{\tau^k}}). \tag{13}$$

5 Experiments

In this section, we present the performance of our RMNet and the other methods to be compared. The following questions are also proposed and investigated:

- **RQ1**: How does our RMNet performs compared to other baselines?
- **RQ2**: How does each module of our RMNet influences the performance?
- **RQ3**: Can our RMNet get balanced feature representation?
- **RQ4**: How is the model complexity of our RMNet compared to other baselines?
- **RQ5**: How effective is our RMNet in an online deployment?

5.1 Experimental Setup

Datasets. We performed experiments on two datasets to demonstrate the effectiveness of our method: MultiTwins [2], a publicly available dataset, and Production, a real-world industrial dataset. Below, we provide detailed information about these datasets:

MultiTwins[1]: This dataset consists of all births in the USA between 1989 and 1991 [2]. And it can be used for multi-treatment effect estimation. For more detailed dataset description, please refer to [19].

Production: This dataset was obtained from one of China's biggest short video platforms. For such platform, video sharpening is known as a valuable source for studying user experience indicators. Different degrees of video sharpening can significantly influence user experiences, potentially affecting users' playback time. To investigate this, we conducted random experiments over two-week, where we assigned three levels of video sharpening ($T = 1, 2, 3$) as treatment groups, while regular videos ($T = 0$) served as the control group. By tracking users' short video playback time during this period, we quantified the impact of definition degradation on user experience. The resulting dataset consists of over 7 million users, with 108 features capturing various user-related characteristics.

Baselines. To evaluate the effectiveness of our RMNet, we compare its performance against several representative methods in neural network-based uplift modeling, they are *S-Learner* [7], *T-Learner* [7], *CFRNet* [13], *MEMENTO* [9] and *HydraNet* [9].

Evaluation Metrics. Following the methodology employed in prior work [3], we utilize two commonly used evaluation metrics in uplift modeling: the Qini coefficient and Kendall's uplift rank correlation. For multi-valued treatment settings, we utilize mQini and sdQini, which correspond to the mean and standard error of the Qini coefficient for all possible treatments, respectively. Similarly, mKendall and sdKendall can be calculated as measures of Kendall's uplift rank correlation.

Implementation Details. All baselines and our RMNet are implemented using PyTorch 1.10. The AdamW optimizer is employed, and the maximum number of training epochs is set to 20. The Optuna package[2] is utilized for parameter search to determine the optimal parameters for all baselines. In our RMNet, the objective of the parameter-tuning process is mQini.

5.2 Overall Performance (RQ1)

In Table 1, we present the comparison results for the MultiTwins and Production datasets, followed by the following observations:

For the MultiTwins dataset: 1) T-Learner exhibits superior performance compared to S-Learner. Despite the utilization of more complex architectures by certain baselines, T-Learner remains competitive in terms of certain metrics. This suggests that the task of uplift modeling in online marketing scenarios differs from traditional individual treatment effect (ITE) estimation. When designing more complex architectures, it is crucial to address essential factors

[1] http://www.nber.org/data/linked-birth-infant-death-data-vital-statistics-data.html.

[2] https://optuna.org/.

that contribute to significant performance improvements, rather than simply extending the model structure without careful consideration. 2) TARNet, CFRNet, MEMENTO, and HydraNet achieve better performance by employing more complex architectures, although the improvement may not be significant across all metrics. This implies that naively extending the structure of binary treatment models without considering effectiveness and efficiency may not yield substantial gains. 3) In contrast to other baselines, our RMNet consistently outperforms all of them in most cases. By utilizing Qini as the objective for hyperparameter tuning, we significantly enhance Qini, even though other metrics may exhibit fluctuations. For the Production dataset: 1) S-Learner and T-Learner remain relatively stable and demonstrate competitive results. 2) Naively extending the binary treatment architecture in the baselines encounters a performance bottleneck, as the abundance of model parameters may lead to learning shocks and make optimization challenging when dealing with numerous treatment groups. 3) Our RMNet considers the effectiveness and efficiency of the uplift modeling structure. On the Production dataset, our RMNet consistently outperforms all baselines, further validating the rationality of the proposed architecture. Through experiments conducted on both a public dataset and a production dataset, we verify the effectiveness of our RMNet. The results demonstrate that our RMNet's architecture achieves superior performance in multi-valued treatment uplift modeling.

5.3 Ablation Study (RQ2)

Furthermore, we conducted ablation studies on our RMNet, aiming to analyze the individual contributions of each module. We systematically removed two components of the RMNet, namely the adversarial training for invariant feature representation (IFR) and the shared reparameterization head (SRH). We created three variants of RMNet, denoted as RMNet (w/o IFR), RMNet (w/o SRH-OR), and RMNet (w/o SRH). Specifically, RMNet (w/o SRH-OR) indicates the removal of orthogonal regularization between the control response head and other uplift prediction heads. The results are presented in Table 2. It is evident from the results that removing any component leads to a degradation in performance. This confirms the effectiveness of each component designed in our RMNet. In particular, the disentangled feature representation module enables selective balancing of feature representation, enhancing the reliability of decision-making. The shared reparameterization module, with a more complex architecture compared to other heads, improves model efficiency and alleviates the inductive bias in uplift prediction. Orthogonal regularization reduces redundancy and enhances the model's performance. All the modules contribute to enhancing the effectiveness of uplift modeling with multi-valued treatment.

5.4 Balance Representation (RQ3)

To evaluate whether our RMNet has indeed learned treatment invariant representations,, we present the visualization in Fig. 3. We depict the T-SNE embeddings

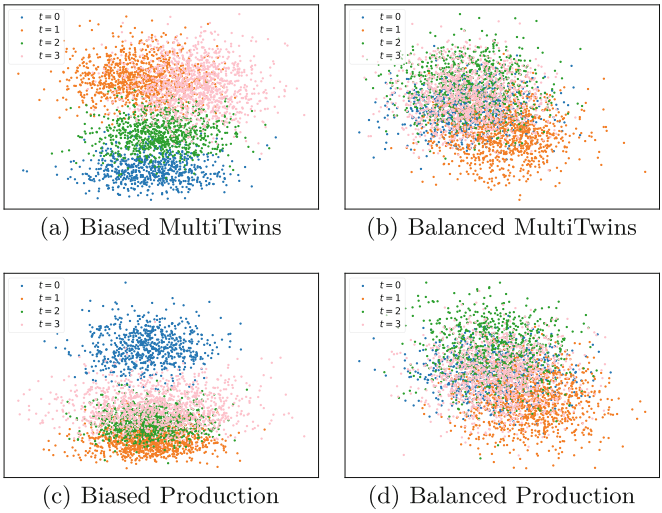
Table 1. Overall comparison between our RMNet and the baselines on MultiTwins and Production datasets, where the best and second best results are marked in bold and underlined, respectively. Note the reported results are the mean ± standard deviation over five runs with different seeds.:

Methods	MultiTwins Dataset				Production Dataset			
	mQini↑	sdQini↓	mKendall↑	sdKendall↓	mQini↑	sdQini↓	mKendall↑	sdKendall↓
S-Learner	0.5930 ± 0.0243	0.1833 ± 0.0466	0.5033 ± 0.0322	<u>0.1814</u> ± 0.0271	0.5807 ± 0.0397	0.2025 ± 0.0377	0.5833 ± 0.0176	<u>0.2133</u> ± 0.0156
T-Learner	0.6586 ± 0.0822	<u>0.1586</u> ± 0.0455	0.6154 ± 0.0814	0.2337 ± 0.0533	0.6021 ± 0.0478	0.2142 ± 0.0353	0.6024 ± 0.0201	0.3089 ± 0.0145
TARNet	0.6659 ± 0.0812	0.2102 ± 0.0854	0.7068 ± 0.0985	0.2065 ± 0.0798	0.6071 ± 0.0544	0.3021 ± 0.0343	0.5799 ± 0.0215	0.3077 ± 0.0197
$\text{CFRNet}_{\text{MMD}}$	0.7586 ± 0.0786	0.1688 ± 0.0766	<u>0.7782</u> ± 0.0925	0.3867 ± 0.0945	<u>0.6823</u> ± 0.0327	0.2785 ± 0.0373	0.5438 ± 0.0278	0.2887 ± 0.0212
$\text{CFRNet}_{\text{WASS}}$	<u>0.7942</u> ± 0.0773	0.3022 ± 0.0699	0.6973 ± 0.0622	0.1897 ± 0.0523	0.6459 ± 0.0386	0.3227 ± 0.0355	0.5996 ± 0.0231	0.3455 ± 0.0182
MEMENTO	0.7374 ± 0.0721	0.2989 ± 0.0642	0.7305 ± 0.0587	0.3575 ± 0.0844	0.6552 ± 0.0322	<u>0.1937</u> ± 0.0284	<u>0.6430</u> ± 0.0231	0.2544 ± 0.0155
HydraNet	0.6908 ± 0.1120	0.2096 ± 0.0839	0.8277 ± 0.0742	0.2587 ± 0.0918	0.6088 ± 0.0321	0.2912 ± 0.0401	0.6202 ± 0.0133	0.2517 ± 0.0221
RMNet	**0.8902** ± 0.0531	**0.1184** ± 0.0594	**0.8304** ± 0.0847	**0.1376** ± 0.0821	**0.7097** ± 0.0322	**0.1877** ± 0.0232	**0.7042** ± 0.0283	**0.2011** ± 0.0137

Table 2. Results of the ablation studies on the Production dataset, where the best and second best results are marked in bold and underlined, respectively. Note the reported results are the mean ± standard deviation over five runs with different seeds.

Methods	mQini↑	sdQini↓	mKendall↑	sdKendall↓
RMNet (w/o IFR)	<u>0.6103</u> ± 0.0304	<u>0.1903</u> ± 0.0192	0.6704 ± 0.0199	0.2677 ± 0.0241
RMNet (w/o SR-OR)	0.6022 ± 0.0204	0.2071 ± 0.0298	0.6255 ± 0.0201	<u>0.2401</u> ± 0.0259
RMNet (w/o SR)	0.5022 ± 0.0213	0.2402 ± 0.0244	<u>0.6887</u> ± 0.0286	0.3018 ± 0.0204
RMNet	**0.7097** ± 0.0322	**0.1877** ± 0.0232	**0.7042** ± 0.0283	**0.2011** ± 0.0137

of both biased and balanced data from both datasets. Our observation reveals that our RMNet successfully mitigates the feature disparity between biased and balanced data, indicating that the learned feature representation Φ achieves a balanced distribution across various treatments and facilitates the model in making more reliable decisions.



(a) Biased MultiTwins (b) Balanced MultiTwins

(c) Biased Production (d) Balanced Production

Fig. 3. T-SNE visualization of the biased data and the learned balanced representation in MultiTwins and Production.

5.5 Complexity Evaluation (RQ4)

We evaluate the model complexity of our RMNet in comparison to other baselines. To ensure a fair comparison among methods with the feature representation module, we maintain identical network structures for this module across TARNet, CFRNet, MEMENTO, HydraNet, and RMNet. The model size and run time are presented in Fig. 4. In particular, the run time is measured over 20

training epochs on the Production dataset, excluding the evaluation time cost. All experiments are conducted on NVIDIA V100 GPUs.

As expected, due to the incorporation of complex shared heads and lightweight heads, the model size of our RMNet is smaller than that of MEMENTO and HydraNet, while remaining competitive with TARNet and CFRNet. Regarding the run time, our RMNet outperforms MEMENTO, HydraNet, and CFRNet. Notably, $\text{CFRNet}_{\text{WASS}}$ exhibits the longest run time among all methods, primarily due to the high computational cost of Wasserstein distance [13]. HydraNet follows with the second longest run time, attributed to the inclusion of a complex target regularization loss. Considering the performance of these methods as shown in Table 1, our RMNet demonstrates an optimal balance of model size, run time, and performance, making it an effective and efficient uplift model.

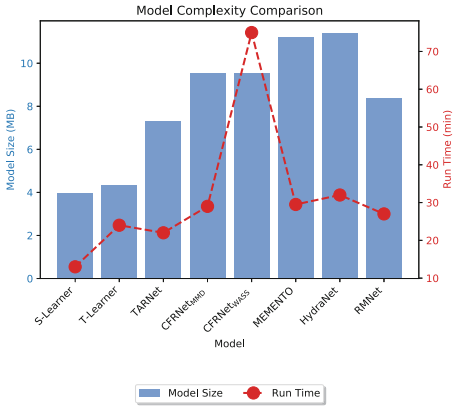


Fig. 4. Model complexity (*i.e.* model size and run time) comparison of all the baselines and our RMNet.

5.6 Online Experiment Results

To test the improvement obtained by using RMNet, we conducted a two-week online A/B experiment. In this experiment, the marketing objective was to enhance users' playback time within the platform by employing different levels of video sharpening as treatments. Specifically, we considered four distinct levels of video sharpening as treatments. We partitioned the online user base into two non-overlapping sets, ensuring that they do not interfere with each other. Each set comprised hundreds of millions of users. The existing method employed on the online platform was a multi-valued treatment Causal Forest (CF) [18]. To evaluate the performance, we adopted ROI (Return On Investment) as the metric. Notably, our RMNet exhibited a 0.5305% improvement over the CF.

6 Conclusion

This paper proposes RMNet, an effective and efficient uplift model for multi-valued treatment. RMNet comprises two key components: 1) an invariant feature representation module that leverages gradient constraints to obtain a balanced representation of all treatments, mitigating selection bias and enhancing model effectiveness for more accurate uplift prediction. 2) a reparameterization multi-head module that treats responses to different treatments as offsets relative to the control response. It employs a reparameterization multi-head structure to effectively reduce the number of model parameters. We also conduct experiments on two datasets to demonstrate the effectiveness and efficiency of RMNet.

Acknowledgements. We thank the support of the National Natural Science Foundation of China (No. 62302310).

References

1. Acharki, N., Lugo, R., Bertoncello, A., Garnier, J.: Comparison of meta-learners for estimating multi-valued treatment heterogeneous effects. In: International Conference on Machine Learning (2023)
2. Almond, D., Chay, K.Y., Lee, D.S.: The costs of low birth weight. Q. J. Econ. **120**(3), 1031–1083 (2005)
3. Belbahri, M., Murua, A., Gandouet, O., Partovi Nia, V.: QINI-based uplift regression. Ann. Appl. Stat. **15**(3), 1247–1272 (2021)
4. Diemert, E., Betlei, A., Renaudin, C., Amini, M.R., Gregoir, T., Rahier, T.: A large scale benchmark for individual treatment effect prediction and uplift modeling. arXiv preprint arXiv:2111.10106 (2021)
5. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., Lempitsky, V.: Domain-adversarial training of neural networks. J. Mach. Learn. Res. **17**(1), 2030–2096 (2016)
6. Kawanaka, S., Moriwaki, D.: Uplift modeling for location-based online advertising. In: Proceedings of the 3rd ACM SIGSPATIAL Workshop, pp. 1–4 (2019)
7. Künzel, S.R., Sekhon, J.S., Bickel, P.J., Yu, B.: Metalearners for estimating heterogeneous treatment effects using machine learning. Proc. Natl. Acad. Sci. **116**(10), 4156–4165 (2019)
8. Liu, D., Tang, X., Gao, H., Lyu, F., He, X.: Explicit feature interaction-aware uplift network for online marketing. arXiv preprint arXiv:2306.00315 (2023)
9. Mondal, A., Majumder, A., Chaoji, V.: Memento: neural model for estimating individual treatment effects for multiple treatments. In: Proceedings of the 31st ACM International CIKM, pp. 3381–3390 (2022)
10. Nie, X., Wager, S.: Quasi-oracle estimation of heterogeneous treatment effects. Biometrika **108**(2), 299–319 (2021)
11. Rubin, D.B.: Causal inference using potential outcomes: design, modeling, decisions. J. Am. Stat. Assoc. **100**(469), 322–331 (2005)
12. Rzepakowski, P., Jaroszewicz, S.: Decision trees for uplift modeling. In: 2010 IEEE International Conference on Data Mining, pp. 441–450. IEEE (2010)
13. Shalit, U., Johansson, F.D., Sontag, D.: Estimating individual treatment effect: generalization bounds and algorithms. In: ICML, pp. 3076–3085. PMLR (2017)

14. Shi, C., Blei, D., Veitch, V.: Adapting neural networks for the estimation of treatment effects. In: Advances in Neural Information Processing Systems, vol. 32 (2019)
15. Sun, Z., Chen, X.: M^3tn: multi-gate mixture-of-experts based multi-valued treatment network for uplift modeling. arXiv preprint arXiv:2401.14426 (2024)
16. Sun, Z., et al.: Robustness-enhanced uplift modeling with adversarial feature desensitization. In: 2023 IEEE International Conference on Data Mining (ICDM), pp. 1325–1330. IEEE (2023)
17. Velasco, B., Cerquides, J., Arcos, J.L.: HydraNet: a neural network for the estimation of multi-valued treatment effects. In: NeurIPS 2022 Workshop on Causality for Real-World Impact (2022)
18. Wager, S., Athey, S.: Estimation and inference of heterogeneous treatment effects using random forests. J. Am. Stat. Assoc. **113**(523), 1228–1242 (2018)
19. Yoon, J., Jordon, J., Van Der Schaar, M.: GANITE: estimation of individualized treatment effects using generative adversarial nets. In: ICLR (2018)
20. Zeng, S., Bayir, M.A., Pfeiffer III, J.J., Charles, D., Kiciman, E.: Causal transfer random forest: Combining logged data and randomized experiments for robust prediction. In: ICDM, pp. 211–219 (2021)
21. Zhao, Z., Harinen, T.: Uplift modeling for multiple treatments with cost optimization. In: 2019 IEEE International Conference on Data Science, pp. 422–431. IEEE (2019)