UTILIZING WORLD MODELS FOR ADAPTIVELY COVARIATE ACQUISITION UNDER LIMITED BUDGET FOR CAUSAL DECISION MAKING PROBLEM

Anonymous authors
Paper under double-blind review

ABSTRACT

Treatment effect estimation from observational data faces critical challenges when covariates are partially observed due to resource constraints or privacy concerns. This study introduces a novel framework leveraging world models (e.g., DeepSeek) to address partial observability in treatment effect estimation using a prompting strategy with few-shot in context learning. Specifically, the world model iteratively prioritizes covariate acquisition based on simulated information gain. It dynamically interacts with historical data and domain knowledge to optimize covariate selection under budget limitations, ensuring efficient data collection for unbiased effect estimation. Experiments on a well-known public available dataset (Twins) show the effectiveness of the proposed framework.

1 INTRODUCTION

Treatment effect estimation lies at the heart of decision-making in fields such as healthcare (Alaa & Van Der Schaar, 2017), economics (Chernozhukov et al., 2013), and public policy (Athey, 2015). In healthcare, for instance, estimating the causal effect of a treatment—such as a drug or surgical procedure—on patient outcomes is critical for clinical guidelines and personalized care. A fundamental challenge in this setting is the requirement to measure all covariates that influence both the treatment assignment and the outcome—to avoid biased estimates. However, in practice, covariates are often partially observed with priority due to resource limitations and privacy concerns, which leads to an important concern: **How can we accurately estimate treatment effects when some covariates are missing, thus making reliable decision?** In addition, we know nothing when a new patient comes to the hospital and we need to perform the tests on the patient in a sequential manner, such as electrocardiogram, blood pressure, etc, which raises another important issue: **how do we prioritize which covariates to collect under constrained budgets?**

World models are used to simulate environments, predict outcomes, and plan actions, which are foundational of intelligent agents. Models like GPT-4 and DeepSeek exhibit remarkable few-shot reasoning abilities, enabling them to generalize from minimal examples and adapt to new tasks. In this paper, we propose to use these capabilities to address the challenge of partial observability in causal inference. Specifically, our framework treats covariate collection as a meta-learning problem: the world model learns, through interaction with historical data and domain-specific knowledge, which covariates are most critical for treatment effect estimation. At each step, the model evaluates the current set of observed covariates, simulates the potential information gain from collecting each missing covariate, and selects the one that maximizes a utility function balancing informativeness and cost. This process continues until the budget is exhausted, at which point treatment effects are estimated using the adaptively collected data.

2 PROBLEM FORMULATION

In our paper, we consider the case that the treatment variable is binary. Suppose we have n data point in total, for each unit i, we collect p covariates, which is denoted as $X_i = (X_{i,1}, \ldots, X_{i,p}) \in \mathbb{R}^p$ and the binary treatment variable is denoted as $T_i \in \{0, 1\}$, where $T_i = 1$ and $T_i = 0$ means assigned and not assigned the treatment, respectively. Let $Y_i \in \mathbb{R}$ be the outcome of interest. For

accurately estimating treatment effect, we adopt the potential outcome framework (Rubin, 1974; Neyman, 1990). Specifically, let $Y_i(0)$ and $Y_i(1)$ be the potential outcome of unit i, where $Y_i(0)$ and $Y_i(1)$ correspond to the outcome for not assigned and assigned treatment, respectively. Since each unit can be assigned with one treatment, thus we can only observe one of $Y_i(0)$ or $Y_i(1)$ corresponding to the treatment value, but not both, which is the well-known fundamental problem of causal inference (Holland, 1986; Morgan & Winship, 2015).

For unit i, the individual treatment effect (ITE) is defined as $\text{ITE}_i = Y_i(1) - Y_i(0)$, which shows that whether the treatment is beneficial for unit i. If $\text{ITE}_i > 0$, we should assign treatment to this unit and vice versa. Meanwhile, the conditional average treatment effect (CATE) is defined as

$$\tau(x) = \mathbb{E}[Y_i(1) - Y_i(0)|X_i = x], \tag{1}$$

which is the expectation of the difference between two potential outcomes given the covariates x. However, as mentioned in the introduction, covariates are often partially observed with priority, we can only achieve the following masked CATE estimation:

$$\hat{\tau}(x_m) = \mathbb{E}[\hat{Y}_i(1) - \hat{Y}_i(0)|X_i = x_i^m], \tag{2}$$

where x_m denotes the masked covariate vector, i.e., $x_i^m = x_i \odot M$ with $M \in \{0, 1\}^p$ and there is exactly m non-zero elements in M, $\hat{Y}_i(1)$ and $\hat{Y}_i(0)$ are the estimated outcome with treatment value equals to 1 and 0 using a world model, respectively. In this paper, we formulate the limited budget scenario by adopting a constraint $m < \gamma$, where γ is a pre-specified hyperparameter. Based on the above analysis, we can formulate the problem as:

$$\min_{x_m} (\hat{\tau}(x_m) - \tau(x))^2, \tag{3}$$

$$\text{s.t. } m \leq \gamma. \tag{4}$$

That is, maximize the estimation accuracy for $\tau(x)$ based on the $\hat{\tau}(x_m)$ with limited budget. In a real-world scenario, we always need to make a decision based on the estimated masked CATE (also known as causal decision making). For example, doctors need to decide if assigning a drug or a surgical procedure to a patient based on the $\hat{\tau}(x_m)$. Thus, we need to minimize the regret below:

$$\text{Regret}(\hat{t}) = \mathbb{E}[Y(t^*(x)) - Y(\hat{t}(x_m))], \tag{5}$$

where $t^*(X)$ is the best optimal treatment assignment based on $\tau(x)$ and $\hat{t}(x_m)$ corresponds to the evaluated treatment assignment policy, such as $\hat{t}(x_m) = \mathbf{1}(\hat{\tau}(x_m) > 0)$. Regret is minimized when $\hat{t}(x_m) = t^*(X)$.

3 METHOD

In this section, we will introduce the few-shot in-context learning method based on a world model. First, note that the decision making problem is usually a Markov decision process, where contains an action space $\mathcal{A}$, a state $\mathcal{S}$, and a transition $\mathcal{P}$. Specifically, the detailed definitions in our scenario are shown below:

• State: s_m. We define the state $s_m = (X, T, Y, X_m)$, where (X, T, Y) is unknown during the evaluation process and invariant over time, and mask vector M controls which covariates are visible.

• Action: a_{m+1}. We define the action as consisting of the selected index of covariates and the stopping criteria at each time step. Specifically, $a_{m+1} \in \{0, 1, \cdots, p\}$ samples from $\pi(X, M)$, which is a $(p + 1)$-dimensional discrete probability distribution. When $a_m \neq 0$ and we still have budget $m < \gamma$, we collect the covariate corresponding to the action, otherwise we stop the acquisition.

• Transition: After choosing a_m as the action, the state $s_{m-1} = (X, W, Y, X_{m-1})$ transitions to $s_m = (X, W, Y, X_m)$, which contains one more covariate compared to the previous state.

Recall that World models can encapsulate the dynamics of how actions affect environments, thus, in this scenario, the world model is used to imagine "*what will happen if we collect this covariate*" and to decide if we can achieve the final goal (minimizing the regret) based on a sequence of actions. Specifically, we use in-context learning with a prompting strategy to induce a high-quality action sequence using the following step:

• **Step 1.** Prompt the world model by providing the background, all variables' name, and some History case, and prompt the world model to imagine the first variable should be collected for a new unit based on the simulated information gain, historical data, and domain knowledge. In addition, explicitly point out that do not use the greedy perspective to imagine.

• **Step 2.** After obtain an output action of the world model, using the current state (the value and name of all collected covariates) as input, prompt the world model to imagine the next action (collect which covariates or stop the acquisition).

• **Step 3.** When reaching the upper bound of the budget or stopping the acquisition, prompt the world model to output the final results.

4 EXPERIMENT

4.1 DATASET

For evaluating the performance, we use the **Twins**[1] (Almond et al., 2005) dataset, which is a seminal resource in causal inference research, collected between 1989 and 1991 through U.S. birth records. It contains observational data on pairs of monozygotic (identical) twins, with the primary aim of studying the causal effect of birth weight on infant mortality and long-term health outcomes. This dataset includes 50 covariates for the twin pair, such as the mother and father's age and education, health complications, and so on. The treatment is defined as the birth weights in grams of both twins in the pair, and the outcome is the mortality outcome for both twins.

4.2 EXPERIMENT DETAILS

To validate whether world models can learn and identify the most critical covariates for decision making through interaction with historical data and domain knowledge, we designed a Progressive Querying experiment based on the state-of-the-art model **Deepseek-R1** (Liu et al., 2024). Specifically, we compare our progressive querying method with the following two baselines: **Random Selection**, **Ask All at Once**, and **Progressive Querying**. The core research question is whether the heavier twin $Y(1)$ has a higher mortality rate than the lighter twin $Y(0)$, i.e., whether $Y(1) > Y(0)$ holds. First, we use GPT-4 to select the most important 20 coariates for answering the core research question, and we pre-define the budget limit is that we can collect 10 covariates at most. The detailed introductions of each strategy are shown below:

- **Random Selection**: Selecting 10 features randomly from 20 background covariates, without considering their importance or relevance. This method represents a completely unstructured approach to feature selection, which may lead to inefficient information acquisition and lower prediction accuracy.
- **Ask All at Once**: The world model selects 10 features at once and makes a decision based on these pre-selected features. This approach follows a static strategy, where the model chooses the 10 features it considers most important based on training data, without acquiring any specific sample information. However, its limitation lies in the lack of adaptability to individual cases, which may result in the omission of critical features.
- **Progressive Querying (Ours)**: As shown in Figure 2, the core idea of this method is to maximize information acquisition and improve prediction accuracy. It leverages historical case analysis and feedback from the current case to iteratively query up to 10 features, dynamically adjusting the selection at each step to ensure that the chosen variables are the most valuable for the final decision. Unlike greedy algorithms, this approach adopts a globally optimal strategy, ensuring that the selected information provides the most comprehensive coverage of decision factors, thereby enhancing prediction accuracy.

By comparing these three approaches, our goal is to verify the effectiveness of the **Progressive Querying** method in causal decision making tasks, specifically in whether it can improve decision accuracy within a limited number of queries.

[1] http://www.nber.org/data/

Background:
We study the causal effect of twin birth weight on mortality rates. Each pair of twins' data includes X (background characteristics) and Y (mortality rate).
(1) X covers parental health status, pregnancy characteristics, medical conditions, and more, with a total of 20 variables.
(2)Y represents the mortality outcome of the twins, where Y0 denotes the mortality of the lighter infant (T0) (0 = survived, 1 = deceased), and Y1 denotes the mortality of the heavier infant (T1) (0 = survived, 1 = deceased).
Task:
Your task is to first analyze the four provided historical cases and observe how background characteristics X influence the relationship between Y0 and Y1. Then, for a new case, you need to progressively select 10 out of the 20 background characteristics for inquiry. The selection strategy should be based on patterns from historical cases to maximize information gain. After collecting 10 characteristics, you must use the existing cases and feature patterns to determine whether Y1 > Y0, meaning whether the mortality rate of the heavier infant is higher than that of the lighter infant.
Interaction Format:
The interaction format is as follows: You play the role of a medical staff member and may ask for one background characteristic at a time. The inquiry strategy should avoid a greedy algorithm and instead consider global optimization. Once 10 characteristics have been collected, you must provide a final judgment: whether the mortality rate of the heavier infant is higher (Y1 > Y0). The background characteristics X include the following 20 variables: mager8 (maternal age), meduc6 (maternal education level), mrace (maternal race), mpre5 (prenatal care initiation time), nprevistq (number of prenatal visits), dfageq (paternal age), feduc6 (paternal education level), frace (paternal race); pregnancy and delivery-related variables: pldel (delivery place), birattnd (attending personnel), adequacy (care adequacy), gestat10 (gestational age, categorized into 10 levels), anemia (anemia), diabetes (diabetes), chyper (chronic hypertension), phyper (pregnancy-related hypertension), eclamp (eclampsia), preterm (history of preterm birth), tobacco (tobacco use), alcohol (alcohol consumption).
History Case:
Below are the complete data for four historical cases:
Case 1: { "pldel": 1, "birattnd": 1, "mager8": 4, "mrace": 1, "meduc6": 3, "mpre5": 1, "adequat": 2, "frace": 1, "gestat10": 5, "anemia": 0, "diabetes": 0, "chyper": 0, "phyper": 0, "eclamp": 0, "preterm": 0, "tobacco": 1, "alcohol": 0, "nprevist": 3, "dfageq": 2, "feduc6": 3, "Y0": 0, "Y1": 1 }.
Case 2: { "pldel": 1, "birattnd": 1, "mager8": 3, "mrace": 2, "meduc6": 3, "mpre5": 1, "adequat": 1, "frace": 2, "gestat10": 4, "anemia": 0, "diabetes": 0, "chyper": 0, "phyper": 0, "eclamp": 0, "preterm": 0, "tobacco": 0, "alcohol": 0, "nprevist": 2, "dfageq": 4, "feduc6": 3, "Y0": 1, "Y1": 0 }.
Case 3: { "pldel": 1, "birattnd": 1, "mager8": 4, "mrace": 1, "meduc6": 3, "mpre5": 1, "adequat": 1, "frace": 1, "gestat10": 3, "anemia": 0, "diabetes": 0, "chyper": 0, "phyper": 0, "eclamp": 0, "preterm": 0, "tobacco": 0, "alcohol": 0, "nprevist": 1, "dfageq": 4, "feduc6": 3, "Y0": 1, "Y1": 0 }.
Case 4: { "pldel": 1, "birattnd": 1, "mager8": 6, "mrace": 1, "meduc6": 5, "mpre5": 1, "adequat": 1, "frace": 1, "gestat10": 3, "anemia": 0, "diabetes": 0, "chyper": 0, "phyper": 0, "eclamp": 0, "preterm": 0, "tobacco": 0, "alcohol": 0, "nprevist": 0, "dfageq": 6, "feduc6": 4, "Y0": 1, "Y1": 0 }.

Please start by progressively asking for 10 characteristics based on the patterns from historical cases, then provide the final judgment on whether Y1 is greater than Y0. Now, you may ask your first question in the format: "Please provide the value of [a specific characteristic]."

Figure 1: The prompt of progressive querying to make causal decision.

4.3 Experiment Results

As shown in Figure 2, the experimental results highlight the effectiveness of the **Progressive Querying** strategy for causal inference in comparison to other feature selection methods. The **Random Selection** strategy performed poorly, by selecting features arbitrarily, it failed to identify the most relevant covariates for decision-making, leading to inaccurate conclusions. Meanwhile, the **Ask All at Once** strategy, though able to extract some meaningful features, still produced suboptimal results. While it considered a broader set of variables, it lacked the adaptability to refine selections dynamically based on case-specific information. Without the ability to update its knowledge base iteratively, this approach was unable to leverage historical data effectively, ultimately leading to an incorrect decision. The **Progressive Querying** strategy, on the other hand, performed well and correctly estimated the causal effect. This success is because our method allows the world model to learn and identify the most critical covariates for decision-making through interaction with historical data and domain knowledge. By progressively refining the selection of features based on earlier queries, we ensured that only the most relevant and impactful covariates were considered, resulting in a more accurate and reliable decision.

	Random Selection	Ask All at Once	Progressive Querying
Selected features	mpre5, gestat10, anemia, diabetes, chyper, phyper, preterm, tobacco, nprevist, feduc6	mager8, meduc6, mpre5, adequat, gestat10, chyper, phyper, preterm, tobacco, nprevist	gestat10, adequat, tobacco, phyper, nprevist, anemia, diabetes, preterm, meduc6, feduc6
Result	Wrong	Wrong	Right

Figure 2: A case study on world model for different feature selection strategies.

In conclusion, the experiment shows that **Progressive Querying** significantly outperforms both the random selection and all-at-once strategies by dynamically adapting to the data and prioritizing the most important features, thus leading to a more accurate causal estimation.

5 Conclusion

This work addresses the challenges of treatment effect estimation under partial covariate observability and budget-constrained data collection. By integrating world models with causal decision making, we propose a dynamic framework that mimics human-like prioritization of covariate acquisition, which uses the model's ability to simulate counterfactual information gain for adaptively selecting covariates. Specifically, we formulate covariate collection as a sequential optimization task guided by domain-specific knowledge and leverage few-shot in context learning method with prompting strategy to achieve more accurate treatment effect estimation. This approach not only enhances the reliability of causal estimates in resource-constrained settings but also aligns with real-world clinical workflows where diagnostic tests are ordered incrementally. Future work should validate the framework on large-scale medical datasets and extend it to settings with heterogeneous costs or temporal dependencies.

References

Ahmed M Alaa and Mihaela Van Der Schaar. Bayesian inference of individualized treatment effects using multi-task gaussian processes. *Advances in neural information processing systems*, 30, 2017.

Douglas Almond, Kenneth Y Chay, and David S Lee. The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083, 2005.

Susan Athey. Machine learning and causal inference for policy evaluation. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 5–6, 2015.

Victor Chernozhukov, Iván Fernández-Val, and Blaise Melly. Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268, 2013.

Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81:945–960, 1986.

Aixin Liu, Bei Feng, Bing Xue, Bingxuan Wang, Bochao Wu, Chengda Lu, Chenggang Zhao, Chengqi Deng, Chenyu Zhang, Chong Ruan, et al. Deepseek-v3 technical report. *arXiv preprint arXiv:2412.19437*, 2024.

Stephen L. Morgan and Christopher Winship. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge University Press, second edition, 2015.

Jerzy Splawa Neyman. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 5:465–472, 1990.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational psychology*, 66:688–701, 1974.