Decision-Focused Sequential Experimental Design: A Directional Uncertainty-Guided Approach

Beichen Wan[1] **Mo Liu**[1] **Paul Grigas**[2] **Zuo-Jun Max Shen**[2]

[1]Department of Statistics and Operations Research, University of North Carolina, Chapel Hill
[2]Department of Industrial Engineering and Operations Research, University of California, Berkeley

[1]{bcwan, mo_liu}@unc.edu [2]{pgrigas, maxshen}@berkeley.edu

Abstract

Classical experimental design has traditionally focused on constructing design variables that facilitate the selection of models with high predictive accuracy. In decision-focused learning, however, the model that achieves the lowest prediction error may not coincide with the one that induces the best downstream decision. Motivated by this misalignment, we investigate appropriate criteria for sequential experimental design in decision-focused settings. Specifically, we consider a sequential data acquisition problem in which a learner adaptively selects samples to label, or equivalently, treatment responses to observe. Existing experimental design methods are inherently decision-blind: they aim to reduce predictive uncertainty, even though reductions in predictive error need not translate into improvements in decision quality. To bridge this gap, we introduce a *directional uncertainty* criterion that aligns predictive uncertainty with the structure of the downstream decision-making problem, in contrast to naïve decision-blind uncertainty quantification methods. We show that this transformation admits strong theoretical guarantees and achieves reduced sample complexity relative to decision-blind design.

1 Introduction

Experimental design seeks to identify, with as little labeled data as possible, a model that is adequate for the task at hand. Here, *data* refers to pairs of (group, response) or, equivalently, (feature, label). This problem is especially compelling when labels are costly or slow to obtain, such as in healthcare studies or when relying on human annotators to evaluate prompts for large language models (LLMs).

Classical sequential experimental design approaches typically aim to minimize prediction error by reducing predictive uncertainty[1, 2, 3, 4]. In decision-focused learning[5, 6, 7, 8], however, predictions are intermediate quantities that feed into a downstream optimization problem, such as pricing, inventory control, or more general linear optimization. What ultimately matters is the *decision loss*, not prediction error alone. Uniformly shrinking predictive uncertainty can therefore fail to improve downstream decisions if the remaining uncertainty does not influence the optimizer in consequential directions.

To address this gap, we develop a sequential design framework that explicitly exploits the structure of the downstream linear optimization problem. The central idea is to replace absolute predictive uncertainty with *directional uncertainty* that captures how errors project onto the decision-relevant directions which perturb the optimizer. This lens yields acquisition rules that prioritize labels most consequential for the downstream decision and de-emphasize uncertainties that are decision-irrelevant.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: MLxOR: Mathematical Foundations and Operational Integration of Machine Learning for Uncertainty-Aware Decision-Making.

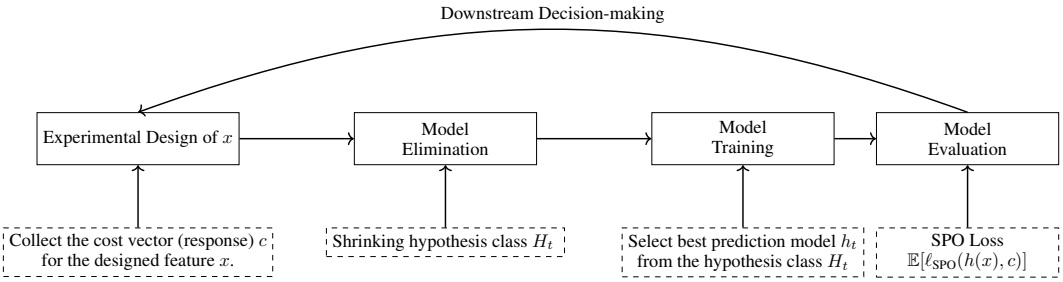


Figure 1: Illustration of Directional Uncertainty-based method and benchmark methods for experimental design.

The process of our sequential design in illustrated in Figure 1. At each iteration, after designing experiment x, we collect the response c, and shrink the hypothesis class according to the new response. Then, the best model is selected from the current hypothesis class. After the experimental design process, the final model is evaluated by the decision-making loss, which is referred to as the SPO (Smart predict-then-optimize) loss.

Motivated by the fact that scaling predictions does not affect downstream solutions in linear optimization, our method prioritizes reducing directional rather than absolute uncertainty. This yields adaptive sampling policies that target labels most consequential for decisions. Our contributions are summarized as follows:

- We propose a sequential design algorithm for decision-focused learning that prioritizes labels by *directional uncertainty*, reflecting the scale-invariance of downstream optimization.
- To ensure convergence, we develop a hypothesis-augmentation method and establish non-asymptotic guarantees, including excess risk bounds and sample complexity, showing our approach is never worse than prediction-uncertainty designs.
- Under a suboptimality gap assumption—meaning that every suboptimal model incurs at least a fixed margin of additional SPO risk relative to the optimal one—we prove our method requires fewer designs than decision-blind baselines to achieve the same decision quality.
- Numerical experiments—on prompt engineering for LLMs and cost-sensitive multi-class classification—demonstrate that our policy achieves comparable decision risk with substantially fewer labels.

1.1 Decision-focused Sequential Design

We study a setting where a decision maker (DM) seeks to learn a predictive model for unknown parameters in a linear optimization problem with cost vector c. The DM adaptively designs experiments, aiming to learn a predictor using few samples while minimizing decision risk.

Let $\|\cdot\|$ denote the ℓ_2 norm, $\mathcal{X}$ the sample space, $\mathcal{C}$ the label space, and $\mathcal{S}$ the polyhedral feasible region. We assume $\mathcal{X}$ has finite support with known marginal $\mu(x)$, each conditional distribution of c given $x = \mathcal{X}_{(i)}$ is bounded, and define radius to be $\rho(\mathcal{C}) := \max_{c\in\mathcal{C}} \|c\|$. We also assume the feasible region is bounded, i.e., $\Delta(\mathcal{S}) := \sup_{w_1,w_2\in\mathcal{S}} \|w_1 - w_2\|$. The DM has no prior knowledge of the joint distribution and must learn $\hat{c} = h(x)$ from a finite hypothesis class $\mathcal{H}$. Let $\mathcal{H}^*$ denote the set of minimizers of SPO risk, and $h^* \in \mathcal{H}^*$ an optimal predictor.

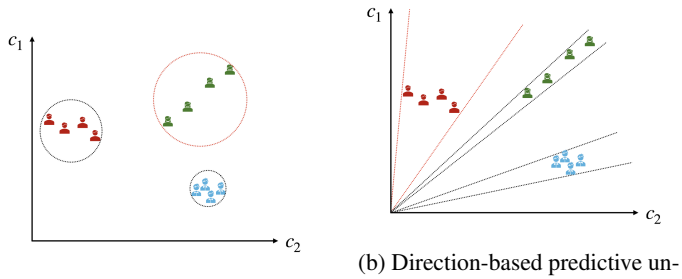
At each period $t = 0, 1, 2, ...$, the DM selects $x_t \in \mathcal{X}$ and observes c_t. Decision loss is measured by the SPO loss [5],

$$\ell_{SPO}(h(x), c) = c^\top w^*(h(x)) - c^\top w^*(c),$$

where $w^*(\hat{c}) = \arg\min_{w\in\mathcal{S}} \hat{c}^\top w$, assuming unique solutions under a consistent tie-breaking rule. The corresponding risk is $R_{SPO}(h) = \mathbb{E}_{(x,c)}[\ell_{SPO}(h(x), c)]$, and the learning objective is $\min_{h\in\mathcal{H}} R_{SPO}(h)$. We define the uncertainty set of predicted labels as $\hat{\mathcal{C}} := \{h(x) : x \in \mathcal{X},\, h \in \mathcal{H}\}$, and the maximum SPO loss by $\omega_{\ell_{SPO}}(\hat{\mathcal{C}}, \mathcal{C}) := \sup_{\hat{c}\in\hat{\mathcal{C}},\, c\in\mathcal{C}} \ell_{SPO}(\hat{c}, c)$.

2 Algorithm and Performance Guarantees

In sequential experimental design, policies often query the covariate with the highest prediction uncertainty rather than those already well learned. A common measure is the ℓ_2 distance among predictions from the current hypothesis class (Figure 2a).



(a) ℓ_2 based predictive uncertainty

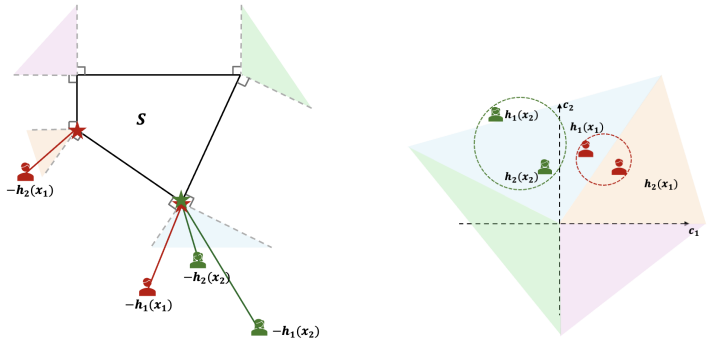
(b) Direction-based predictive uncertainty

Figure 2: Comparison of two predictive uncertainty measures.

Figure 2a illustrates selecting one of three data points with the greatest uncertainty under four predictors. The target cost vector $c \in \mathbb{R}^2$ is two-dimensional; red, green, and blue denote x_1, x_2, x_3 evaluated by $h_1, \ldots, h_4 \in \mathcal{H}$. The circle marks the maximum ℓ_2 distance among their predictions, and hence the green point is chosen.

However, an observation shows that SPO loss is scale invariant: For all $\hat{c}$ and $\alpha > 0$, $\ell_{SPO}(\alpha\hat{c}, c) = \ell_{SPO}(\hat{c}, c)$ since $w^*(\hat{c}) = w^*(\alpha\hat{c})$. This indicates that for a fixed true cost vector c, $\ell_{SPO}(., c)$ is scale independent for any given prediction $\hat{c}$. This comes from the geometric property of linear programming; the oracle $w^*(\hat{c})$ only depends on the estimated cost vector's direction, not its length. The observation also shows that the ℓ_2 based criterion is decision-blind, and hence quantifying prediction uncertainty using the ℓ_2 distance may lead to inefficiency in decision-focused problems. Example 1 further illustrates this idea.

Thus, it motivates us to expand beyond the canonical distance metrics to construct a more effective, decision-focused sequential experimental design algorithm. A natural extension is to utilize directional information to evaluate the predictive uncertainty, illustrated in Figure 2b.



(a) Cost vector prediction with corresponding optimizer

(b) Cost vector prediction with ℓ_2 distance

Example 1 *Consider a two-dimensional setting with hypothesis class $\mathcal{H} = h_1, h_2$ and data points $\mathcal{X} = x_1, x_2$. We choose one point to label and denote predictions by $h_i(x_j)$. Figure 3a shows the feasible region $\mathcal{S}$, predicted cost vectors (plotted in the negative direction), and their optimizers. Each cost vector lies in a normal cone, with the corresponding vertex as optimizer. In Figure 3b, although ℓ_2 uncertainty is larger at x_2 than x_1, predictions at x_2 lead to the same optimizer, while predictions at x_1 yield different optimizers. Thus, from the perspective of the SPO loss, $\ell_{\text{SPO}}(h_1(x_1)) \neq \ell_{\text{SPO}}(h_2(x_1))$ but $\ell_{\text{SPO}}(h_1(x_2)) = \ell_{\text{SPO}}(h_2(x_2))$, despite x_2 exhibiting greater ℓ_2 distance.*

2.1 Algorithm: Importance Weighted Sequential Experimental Design

Our importance-weighted sequential design algorithm based on prediction uncertainty (IWSD-PU) is given in Algorithm 1. At each iteration t, the algorithm selects a design x_t from the pool $\mathcal{X} = \{\mathcal{X}_{(i)}\}_{i=1}^m$. It maintains a confidence set $H_t \subseteq \mathcal{H}$, initialized as $\mathcal{H}$. To compute sampling probabilities, we evaluate $p_{t,i} = \max_{h_1,h_2 \in H_t} \left\| \frac{h_1(x_i)}{\|h_1(x_i)\|} - \frac{h_2(x_i)}{\|h_2(x_i)\|} \right\|$, and set $\pi_{t,i} = \frac{p_{t,i}}{\sum_{j=1}^m p_{t,j}}$. The design x_t is drawn according to $(\pi_{t,1}, \ldots, \pi_{t,m})$, with $Q_{t,i} = \mathbb{I}\{x_t = \mathcal{X}_{(i)}\}$. After selecting x_t, the algorithm observes $c_t \sim p(c \mid x_t)$ and adds (x_t, c_t) with weight $\pi_{t,i}$ to the training set W_{t-1}.

Pseudo Algorithm 1 Importance-Weighted Sequential Design Based on Prediction Uncertainty (IWSD-PU)

Input: Initial slackness r_0.
Set $W_0 \leftarrow \emptyset$, $n_0 \leftarrow 0$, $H_0 \leftarrow \mathcal{H}_{\text{augmented}}$.
while condition for ending loop **do**
  For each potential design $\mathcal{X}_{(i)} \in \mathcal{X}$, calculate $p_{t,i} \leftarrow \max_{h_1,h_2 \in H_t} \left\{ \| \frac{h_1(\mathcal{X}_{(i)})}{\|h_1(\mathcal{X}_{(i)})\|} - \frac{h_2(\mathcal{X}_{(i)})}{\|h_2(\mathcal{X}_{(i)})\|} \| \right\}$, $\pi_{t,i} = \frac{p_{t,i}}{\sum_{j=1}^m p_{t,j}}$
  if $\sum_{j=1}^m p_{t,j} > 0$. **then**
    Sample x_t according to the probability $\pi_t = (\pi_{t,1}, \ldots, \pi_{t,m})$. Let $Q_{t,i} \in \{0,1\}$ denotes the indicator for the realization of x_t.
  else
    stop the algorithm and return h_T
  end if
  Conduct experiment and obtain a label c_t associated with x_t
  Update $W_t \leftarrow W_{t-1} \cup \{(x_t, c_t, \pi_t, Q_t)\}$, $n_t \leftarrow n_{t-1} + 1$.
  Let $\hat{\ell}^t(h) \leftarrow \frac{1}{t} \sum_{(x_j,c_j,\pi_j,Q_j) \in W_t} \frac{\mu(x_j)}{\pi_j} Q_j \ell_{SPO}(h(x_j), c_j)$.
  Update $h_t \leftarrow \arg\min_{h \in H_{t-1}} \hat{\ell}^t(h)$ and $\hat{\ell}^{t,*} \leftarrow \min_{h \in H_{t-1}} \hat{\ell}^t(h)$.
  Update the confidence set of the predictor H_t by $H_t \leftarrow \{h \in H_{t-1} : \hat{\ell}^t(h) \leq \hat{\ell}^{t,*} + r_t\}$.
  Update $r_{t+1} \leftarrow r_t \sqrt{\frac{t}{t+1} \frac{\log(2(t+1))}{\log(2t)}}$
end while
Return h_T.

In Algorithm 1, the random variables at iteration t are (x_t, c_t, π_t, Q_t). For simplicity, we use random variable $z_t \in \mathcal{Z} := \mathcal{X} \times \mathcal{C} \times (0,1]^m \times \{0,1\}^m$ to denote the tuple of random variables $z_t := (x_t, c_t, \pi_t, Q_t)$. Thus, z_t depends on $z_1, ..., z_{t-1}$ and the classical convergence results for i.i.d. samples do not apply to our importance-weighted sampling algorithm. We define $\mathcal{F}_{t-1}$ as the σ-field of all random variables until the end of iteration $t-1$, i.e., $z_1, ..., z_{t-1}$. With slight abuse of notation, in Algorithm 1, the re-weighted loss function at iteration t can be rewritten as $\ell^{\texttt{rew}}(h; (x_t, c_t, \pi_t)) := \sum_{i=1}^m \frac{\mu(x_i)}{\pi_{t,i}} Q_{t,i} \ell_{SPO}(h(x_i), c_i)$

It can be shown that the expectation of the re-weighted loss is an unbiased estimator for the risk of h. Thus, under appropriate regularity conditions that are guaranteed by our assumptions, the empirical re-weighted loss is expected to converge to the risk, and its minimizer h_t over the confidence set $\mathcal{H}_{t-1}$ is expected to converge to the true h^* that minimizes the risk.

Theorem 1 shows the main convergence result of the algorithm, which relies on a mild assumption 1 about the margin condition that can guarantee a Lipschitz-like condition in [9].

Assumption 1 (Directional Margin Condition) *There exists $\tau > 0$ such that for all $h \in \mathcal{H}, x \in \mathcal{X}$, it holds that $\frac{\nu_{\mathcal{S}}(h(x))}{\|h(x)\|} \geq \eta$.*

Theorem 1 (Risk Bound) *Suppose that Assumption 1 holds. Let $\delta \in (0,1]$ be a given parameter, and set $r_0 \leftarrow \max\{2\omega_\ell(\hat{\mathcal{C}}, \mathcal{C}), 2\gamma L \sqrt{ln(2|\mathcal{H}|/\delta)}\}$ where $L = \frac{\Delta(\mathcal{S})\rho(\mathcal{C})}{2\eta}$. Then, the following*

guarantees hold with probability at least $1-\delta$: for all $T \geq 1$, the SPO risk satisfies

$$R_{\text{SPO}}(h_T) - R_{\text{SPO}}(h^*) \leq 4\gamma L\sqrt{\frac{\log(2T|\mathcal{H}|/\delta)}{T}}. \tag{1}$$

Theorem 1 establishes that the excess risk of our algorithm admits a bound of order $\tilde{\mathcal{O}}(1/\sqrt{T})$, matching the standard rate in experimental design. Building on this, Theorem 2 shows that, for a certain class of distributions, our decision-focused design attains strictly lower sample complexity than the decision-blind design. Specifically, we define this class as $\mathcal{P}_\alpha := \mathcal{D} : R_{\text{SPO}}(h) - R^*_{\text{SPO}} \geq \alpha$ for all $h \in \mathcal{H} \setminus \mathcal{H}^*$, which enforces a uniform suboptimality gap in the SPO risk.

Theorem 2 (Earlier stopping time than decision-blind design) *Suppose that Assumption 1 holds. Considering the class of distribution within $\mathcal{P}_\alpha$ for some $\alpha > 0$, we have the following results:*

1. *For Algorithm 1, for any $\delta \in (0,1)$, there exists a constant T_0 (depending on α and δ), such that for any distribution $\mathcal{D} \in \mathcal{P}_\alpha$ that satisfies Assumptions 1, for any $T \geq T_0$, with probability at least $1-\delta$, the candidate prediction set H_T has excluded all suboptimal prediction models, i.e., $H_T \subseteq \mathcal{H}^*$.*
2. *For the decision-blind design, there exists a distribution of $\mathcal{D} \in \mathcal{P}_\alpha$ and a constant δ_0, such that for any T, the probability of having some suboptimal prediction models within H_T is larger than δ_0, i.e., with probability at least δ_0, there exists $h \in H_T$, such that $R_{\text{SPO}}(h) > R^*_{\text{SPO}}$.*

Conclusion. We propose a new design criterion that is based on the directional uncertainty of the response, in the context of decision-focused sequential experimental design. We demonstrate that this simple transformation of the uncertainty enjoys a $\tilde{\mathcal{O}}(1/\sqrt{T})$ convergence rate in the general case, and an exponential convergence rate in Theorem 2. We also empirically check the effectiveness of our direction uncertainty-based design in real-world data.

References

[1] Rui Castro, Rebecca Willett, and Robert Nowak. Faster rates in regression via active learning. In *NIPS*, volume 18, pages 179–186, 2005.

[2] Masashi Sugiyama and Shinichi Nakajima. Pool-based active learning in approximate linear regression. *Machine Learning*, 75(3):249–274, 2009.

[3] Wenbin Cai, Muhan Zhang, and Ya Zhang. Batch mode active learning for regression with expected model change. *IEEE transactions on neural networks and learning systems*, 28(7):1668–1681, 2016.

[4] Alina Beygelzimer, Sanjoy Dasgupta, and John Langford. Importance weighted active learning. In *Proceedings of the 26th annual international conference on machine learning*, pages 49–56, 2009.

[5] Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1):9–26, 2022.

[6] Jayanta Mandi, Peter J Stuckey, Tias Guns, et al. Smart predict-and-optimize for hard combinatorial optimization problems. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 1603–1610, 2020.

[7] Brandon Amos and J Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In *International conference on machine learning*, pages 136–145. PMLR, 2017.

[8] Priya Donti, Brandon Amos, and J Zico Kolter. Task-based end-to-end model learning in stochastic optimization. *Advances in neural information processing systems*, 30, 2017.

[9] Othman El Balghiti, Adam N Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. *Advances in neural information processing systems*, 32, 2019.