Persuasive Prediction via Decision Calibration

Jingwu Tang* [1] Jiahao Zhang* [1] Fei Fang [1] Zhiwei Steven Wu [1]
[1]Carnegie Mellon University

Persuasion without Prior?

Bayesian persuasion studies how a sender, who privately observes a state drawn from a prior distribution, strategically sends a signal to influence a receiver's action and maximize their **own utility**.

Key assumption in BP: Sender and receiver share the precise knowledge of the **prior**.

❗ Exact prior assumption is sometimes unrealistic, especially when the state space is **large** or even infinite.

🔑 Question: Can the sender using only **finite data** achieve utility comparable to that of a fully informed Bayesian sender with **exact** knowledge of the distribution $\mathcal{D}$?

Problem Setting and the Objective

Data: $(x, y) \sim \mathcal{D}$, where $x \in \mathcal{X}, y \in [-1, 1]^d$.

Randomized Predictor: $f \in \Delta(\mathcal{H})$, where $\mathcal{H}$ is a hypothesis class of deterministic predictors $h : \mathcal{X} \to \mathcal{Y}$.

Receiver Model:

1. N receivers, each with actions $\mathcal{A}_i$ ($|\mathcal{A}_i| = m$);
2. Utility: $v_i(a, y)$ (linear, L-Lipschitz in y);
3. Best response: $b_i(h(x), a_i) = \begin{cases} 1 & \text{if } a_i = \arg\min v_i(a_i', h(x)) \\ 0 & \text{otherwise} \end{cases}$

❓ Question: How can we ensure downstream agents trust the prediction and best respond?

Decision Calibration: A predictor f is ϵ-decision calibrated if:

$$\mathrm{DecCE}(f) = \max_{i \in [N], j \in [d], a \in \mathcal{A}_i} \left| \mathbb{E}_{h,f} \mathbb{E}_{\mathcal{D}} [(y_j - h(x)_j) b_i(h(x), a)] \right| \le \epsilon.$$

✅ Agents that best respond to a decision calibrated predictor have low swap regret!

Sender Model: The sender's utility $u(\boldsymbol{a}, y)$ depends on the joint action $\boldsymbol{a}$ and the outcome y (linear in y).

Sender's Objective: provide predictions that ensure agents have no regret when best responding, while maximizing their **own** expected utility:

$$\max_f \mathbb{E}\left[\sum_{\boldsymbol{a}} u(\boldsymbol{a}, y)\, b(h(x), \boldsymbol{a})\right] \quad \text{s.t. } \mathrm{DecCE}(f) \le \gamma.$$

where $b(h(x), \boldsymbol{a}) = \prod_i b_i(h(x), a_i)$.

A Minimax Approach

The algorithm we use follows a minimax approach: we introduce the Lagrangian and reformulate the problem as a minimax game.

$$\min_{f \in \Delta\mathcal{H}} \max_{\lambda \in \mathbb{R}_+^{2Nmd}} \mathcal{L}_{\mathcal{D}}(f, \lambda) := -\mathbb{E}_{h \sim f} \mathbb{E}_{\mathcal{D}} \left[\sum_{\boldsymbol{a} \in \mathcal{A}} u(\boldsymbol{a}, y) \cdot b(h(x), \boldsymbol{a}) \right]$$
$$+ \sum_{s \in \{+,-\}} \sum_{i=1}^{N} \sum_{j=1}^{d} \sum_{a_i \in A_i} \lambda_{s,i,j,a_i} s \Big(\mathbb{E}_f \mathbb{E}_{\mathcal{D}} [(h(x)_j - y_j) \cdot b_i(h(x), a_i)] - \gamma \Big). \tag{1}$$

Instead of considering unbounded $\lambda \in \mathbb{R}_+^{2Nmd}$, we restrict to $\lambda \in \Lambda = \{\lambda' | \lambda' \in \mathbb{R}_+^{2Nmd}, \|\lambda\|_1 \le C\}$.

Now we can solve the minimax game with **Best Response vs. No Regret Dynamics**.

ERM Oracle: $\mathrm{ERM}(D, \lambda) = \arg\min_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^{n} \ell_\lambda(h, x_i, y_i)$ where

$$\ell_\lambda(h, x, y) = -\sum_{\boldsymbol{a} \in \mathcal{A}} u(\boldsymbol{a}, y) \cdot b(h(x), \boldsymbol{a})$$
$$+ \sum_{s \in \{+,-\}} \sum_{i=1}^{N} \sum_{j=1}^{d} \sum_{a_i \in A_i} \lambda_{s,i,j,a_i} s\big([(h(x)_j - y_j) \cdot \tilde{b}_i(h(x), a_i)] - \gamma \big).$$

Algorithm 1:
For each iteration $t = 1, 2, ..., T$:

1. Learner best responds to λ_t: Use the ERM oracle to compute $h_t = \mathrm{ERM}(D, \lambda_t)$.
2. Auditor runs Hedge to obtain λ_{t+1}:

Outputs $\hat{f} = \mathrm{Uniform}(h_1, \cdots, h_T)$.

Theorem 1: Suppose Algorithm 1 runs for $T = O(\log(Nmd)/\epsilon^4)$ rounds with dataset of size $n \ge O(\frac{\log(|\mathcal{H}|Ndm/\delta)}{\epsilon^4})$. With probability at least $1 - \delta$, it outputs $\hat{f}$ that satisfies

1. $\mathrm{DecCE}(\hat{f}) \le \gamma + \epsilon$;
2. If receivers use strict best response:
   * Their swap regret is at most $2mL(\gamma + \epsilon)$;
   * Sender utility is ϵ-optimal.

Challenges with Infinite Hypothesis Classes

Even if two predictors $h_1, h_2 \in \mathcal{H}$ are close (i.e., $\sup_{x \in \mathcal{X}} \|h_1(x) - h_2(x)\|_\infty$ is small), the sender's utility can differ greatly under strict best response!

Example: Let $\mathcal{Y} = [0, 1]$, with $\Pr[y \mid x] = 0.5$ for all x. Receiver chooses a or a', with utility: $v(a, y) = y$, $v(a', y) = 1 - y$. → Best response: a if $y \le 0.5$, a' if $y > 0.5$. Sender's utility: $u(a, y) = 1$, $u(a', y) = 0$. Consider predictors: $h_1(x) = 0.5 - \epsilon$, $h_2(x) = 0.5 + \epsilon$. Both are ϵ-decision calibrated, but: Sender utility: 1 under h_1, 0 under h_2.

Matching the Bayesian Persuasion Benchmark

State space Θ and prior μ: $\mu_{\mathcal{D}}(\theta) = \Pr_{(x,y) \sim \mathcal{D}}[\mathbb{E}[y \mid x] = \theta]$.

Signaling scheme: $\pi : \Theta \to \Delta(S)$.

Receiver with utility funuction v_i selects the best action based on the posterior distribution upon observing the signal.

Calibration: $\mathrm{CE}(f) := \mathbb{E}_{h \sim f, (x,y) \sim \mathcal{D}}[(y - h(x)) | h(x) = v] = 0, \forall v \in \mathcal{Y}$.

Lemma 1: The best response action to any perfectly calibrated predictor corresponds to a valid signaling scheme that persuades the receiver to take that action.

Lemma 2: For any perfectly decision-calibrated predictor, there exists a perfectly calibrated predictor that induces the same best response action of the receiver.

Therefore, given a perfectly decision-calibrated predictor, we can recover a valid signaling scheme that persuades the receiver to take the recommended action.

Theorem 2: $\hat{f}$ achieves the sender's utility of the optimal signaling scheme induced by a perfectly decision-calibrated predictor in $\Delta(\mathcal{H})$.

Persuasive Prediction under Quantal Response

Because best response is **discontinuous**, when $\mathcal{H}$ is infinite, estimating sender utility is hard—even if $\mathcal{H}$ **has low complexity**!

❓ Question: What alternative agent behavior models can we consider when learning a predictor from an infinite hypothesis class?

Quantal Response: The probability that receiver i selects action a_i given the prediction $h(x)$:

$$\tilde{b}_i(h(x), a_i) = \frac{e^{\eta v_i(a_i, h(x))}}{\sum_{a_i' \in \mathcal{A}_i} e^{\eta v_i(a_i', h(x))}}.$$

Smoothed Decision Calibration:

$$\mathrm{SmDecCE}(f) :$$
$$= \max_{i \in [N], j \in [d], a in \mathcal{A}_i} \left| \mathbb{E}_{h \sim f} \mathbb{E}_{(x,y) \sim \mathcal{D}} \left[(y_j - h(x)_j) \cdot \tilde{b}_i(h(x), a) \right] \right|.$$

f is said to be ϵ-decision calibrated if $\mathrm{SmDecCE}(f) \le \epsilon$.

Replacing b with $\tilde{b}$, we design a variant of Algorithm 1 that provides similar guarantees when agents follow the quantal response model.