Federated CP New Setting

Min, Xia

May 28, 2024

1 Idea1

For K agents each with $Z_k = \{(X_i^k, Y_i^k)\}_{i=1}^{n_k}$ sampled from distribution P^k, and point predictor f_1 based score S_i^k. If all $P^k = P$ and a new test point X, Y from P^1. For any trial data y and score S follow the procedure in "Conformal prediction with local weights: randomization enables robust guarantees"[1]:

- Find some kernel function $H(\cdot,\cdot)$, sample $\tilde{X}$ based on $H(X,\cdot)$.
- Calculate empirical function $\tilde{F} = \sum_{i,k} w_i^k \delta_{S_i^k} + w\delta_S$ with weight

$$w_i^k = \frac{H(X_i^k, \tilde{X})}{\sum_{i',k'} H(X_{i'}^{k'}, \tilde{X}) + H(X, \tilde{X})},\ w = \frac{H(X, \tilde{X})}{\sum_{i',k'} H(X_{i'}^{k'}, \tilde{X}) + H(X, \tilde{X})}.$$

- Conformal set is $C_\alpha(X) = \{S \leq Q(1-\alpha, \tilde{F})\}$.

However for all k $P^k = P$ is not practical, potential covariate shift exists

1.1 Experiment1

Assume agent $k = 1, \cdots, K$ each has n samples $X_1^k, \cdots, X_n^k$ follow $N(\mu_k, 9)$. Sythesize $Y_i^k = (X_i^k)^2 + \epsilon$, where $\epsilon \sim N(0, (ep * |X|)^2)$, ep be some parameter. Under this problem only have covariate shift.

- X has different distribution for each agents

- $EY|X$ is same for all agents
- $Y - EY|X$ has same distribution for all agents

Covariate shift has little influence on this method as the method is localized. More on Fig 1.

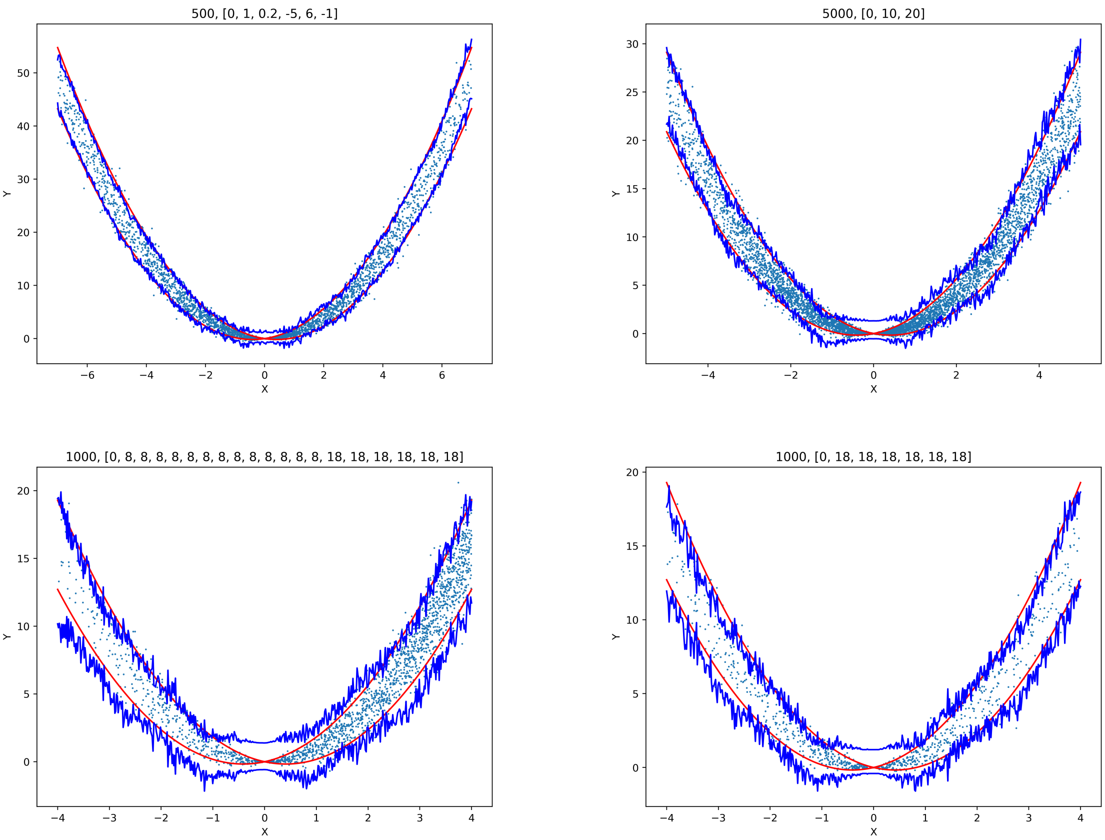

Figure 1: Fig1

Generate 1000 samples with location 0, and 1000 samples with location 0 and 20. The left shows the performance on only within samples with 0 location and 1000 samples is better than the right one using 2000 data samples. This comes from covariate shift, the scores from samples with location 20 influences origin behavior, from Fig2

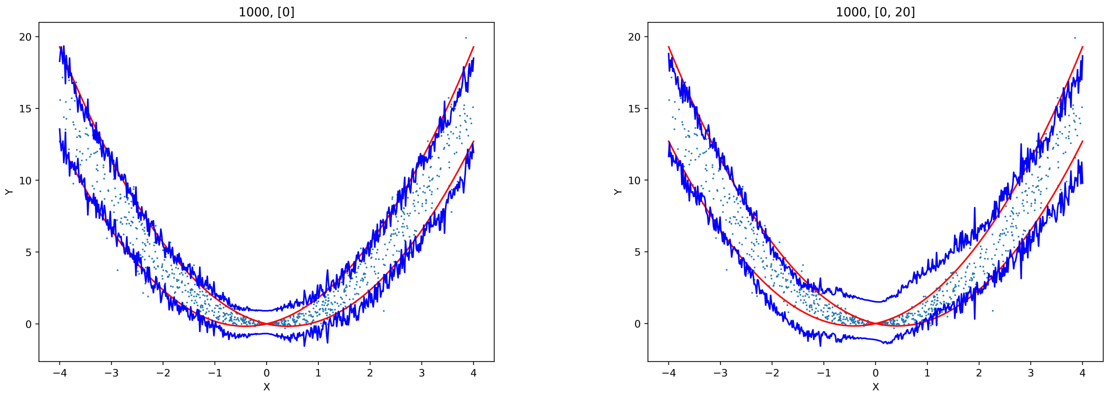

Figure 2: Fig2

1.2 Experiment2

Assume agent $k = 1, \cdots, K$ each has n samples $X_1^k, \cdots, X_n^k$ follow $N(\mu_k, 9)$. Sythesize $Y_i^k = (X_i^k)^2 + \epsilon$, where $\epsilon \sim N(0, (ep * |X| + ep * \theta_k)^2)$, ep be some parameter and θ_k generated for each agent randomly between 1 and 10.

- X has different distribution for each agents
- $EY|X$ is same for all agents
- $Y - EY|X$ has different distribution for different agents

References

[1] Rohan Hore and Rina Foygel Barber. Conformal prediction with local weights: randomization enables local guarantees. *arXiv preprint arXiv:2310.07850*, 2023.