NEURAL INFORMATION PROCESSING SYSTEMS

Boosted Conformal Prediction Intervals

Ran Xie, Rina Foygel Barber, Emmanuel J. Candès

NeurIPS 2024

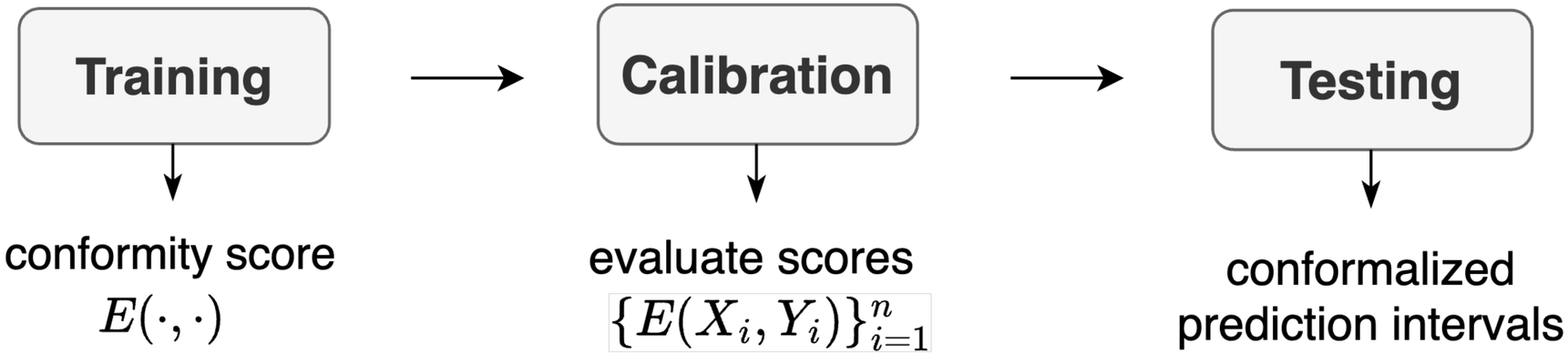
Dec, 2024

Stanford University

THE UNIVERSITY OF CHICAGO

Background: Conformal Prediction

The split conformal procedure



Example $E(x, y) = |y - \hat{f}(x)|$

$$C_n(X_{n+1}) = \hat{f}(X_{n+1}) \pm Q_{1-\alpha}(\{E(X_i, Y_i)\}_{i=1}^n)$$

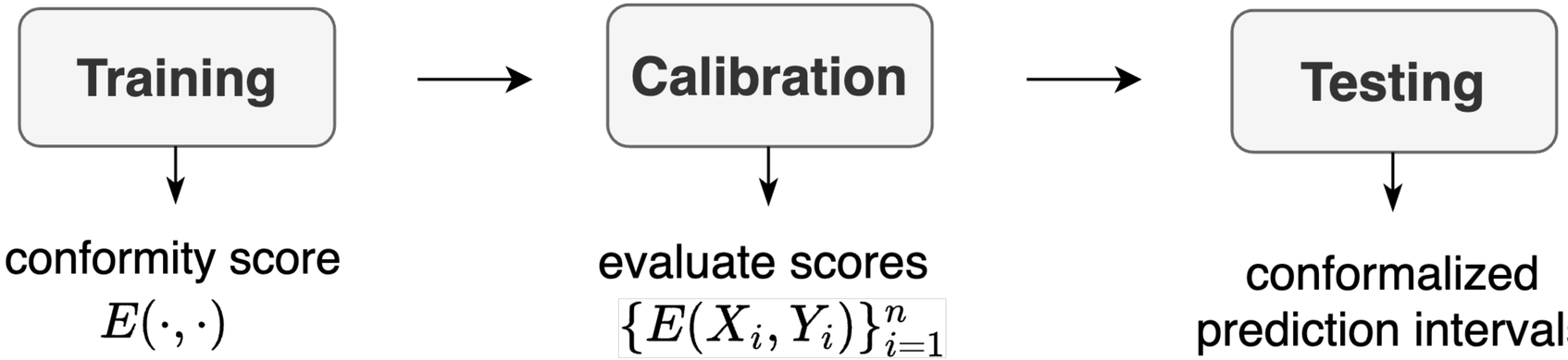
Other conformity scores: Local, CQR, etc.

$$P\left(Y_{n+1} \in C_n(X_{n+1})\right) = 90\,\%$$

Other properties? e.g. Conditional Coverage: $P\left(Y_{n+1} \in C_n(X_{n+1}) \,|\, X_{n+1} = x\right) = 90\,\%$

Background: Conformal Prediction

The split conformal procedure



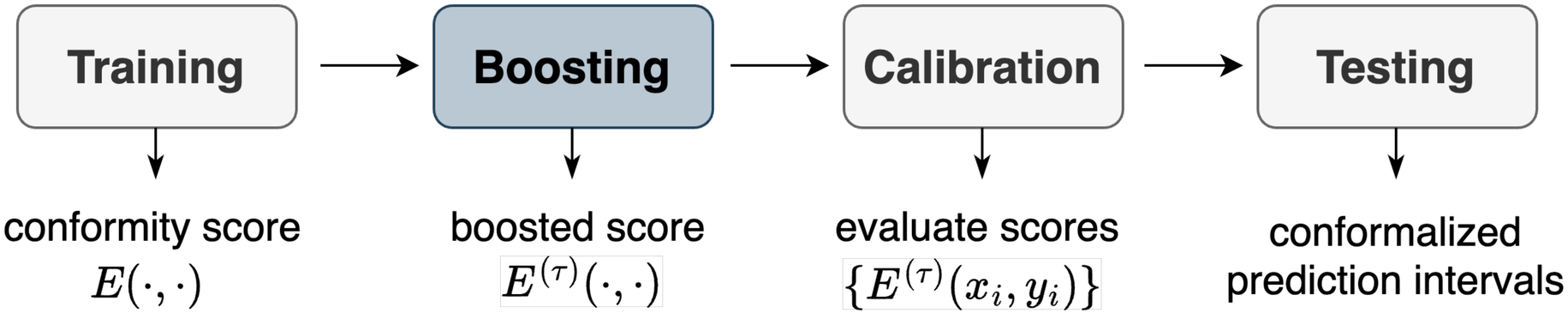
Example $E(x, y) = |y - \hat{f}(x)|$

$$C_n(X_{n+1}) = \hat{f}(X_{n+1}) \pm Q_{1-\alpha}(\{E(X_i, Y_i)\}_{i=1}^n)$$

Question: How should we choose the conformity score function?

Introducing BoostedCP

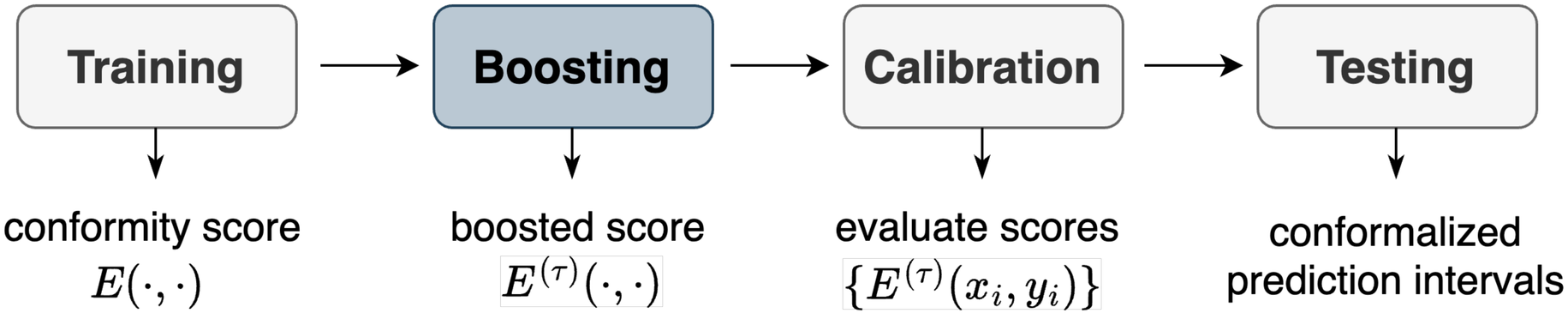
The boosted conformal procedure



- targets **user-specified properties** (e.g. conditional coverage, interval length)
- searches within a family that contains the oracle
- operates **post-training**
- maintains marginal coverage guarantee

Introducing BoostedCP

The boosted conformal procedure



- targets **user-specified properties** (e.g. conditional coverage, interval length)
- searches within a family that contains the oracle
- operates **post-training**
- maintains marginal coverage guarantee

Boosting for conditional coverage

Custom loss via contrast trees

Ideal prediction interval with valid conditional coverage satisfies

$$P\left(Y \in C_n(X_{n+1}) \,|\, X_{n+1} = x\right) \approx 1 - \alpha$$

- Define the absolute within-group deviation for each index group R as

$$d\left(C_n(\,\cdot\,); R\right) = \left| \,|R|^{-1} \sum_{j \in R} 1(Y_j \in C_n(X_j)) - (1 - \alpha) \right|$$

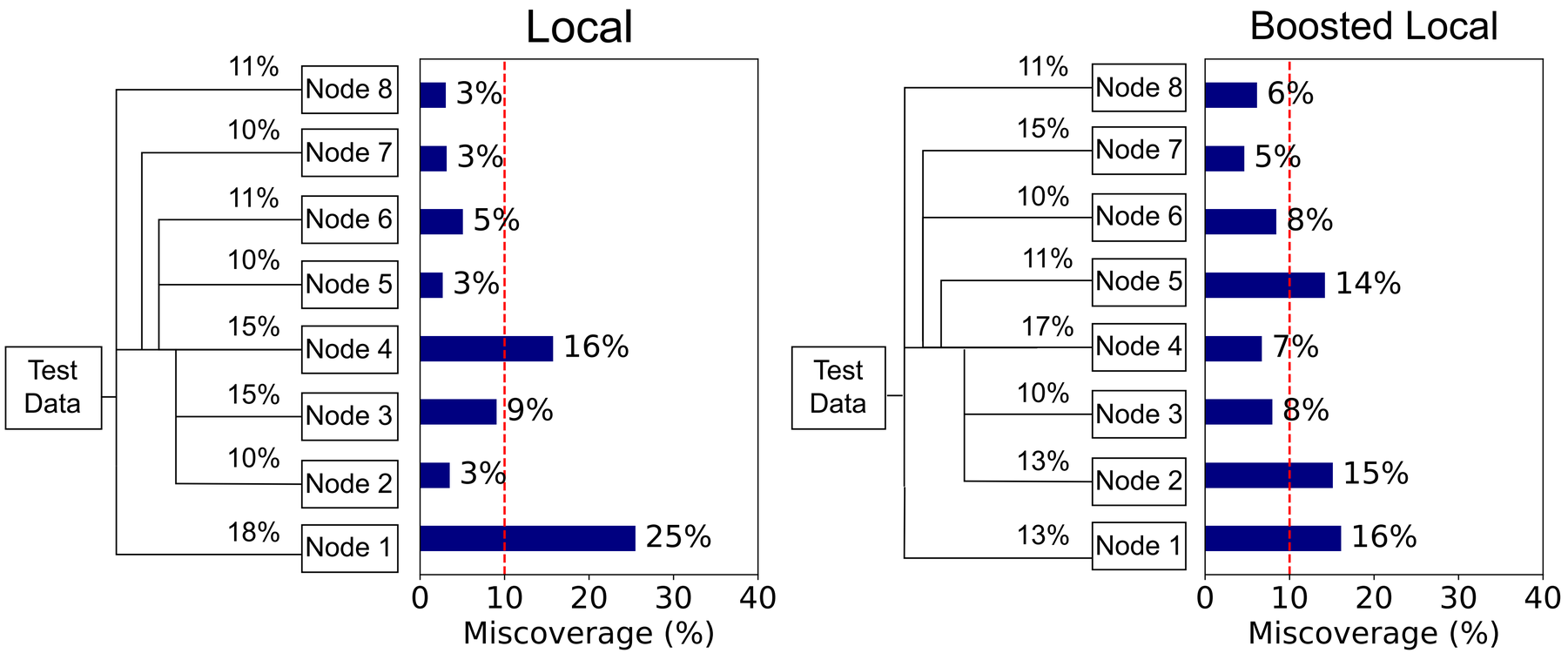
- Use Contrast Trees to iteratively identifies splits within the feature space that maximizes absolute within-group deviation

- Set the conditional coverage loss of the prediction interval as

$$\ell_M(E) = \max_{1 \le m \le M} d\left(C_n(\,\cdot\,); R_m\right)$$

Boosting for conditional coverage

Empirical results



Thank you for your attention!

Paper

Code