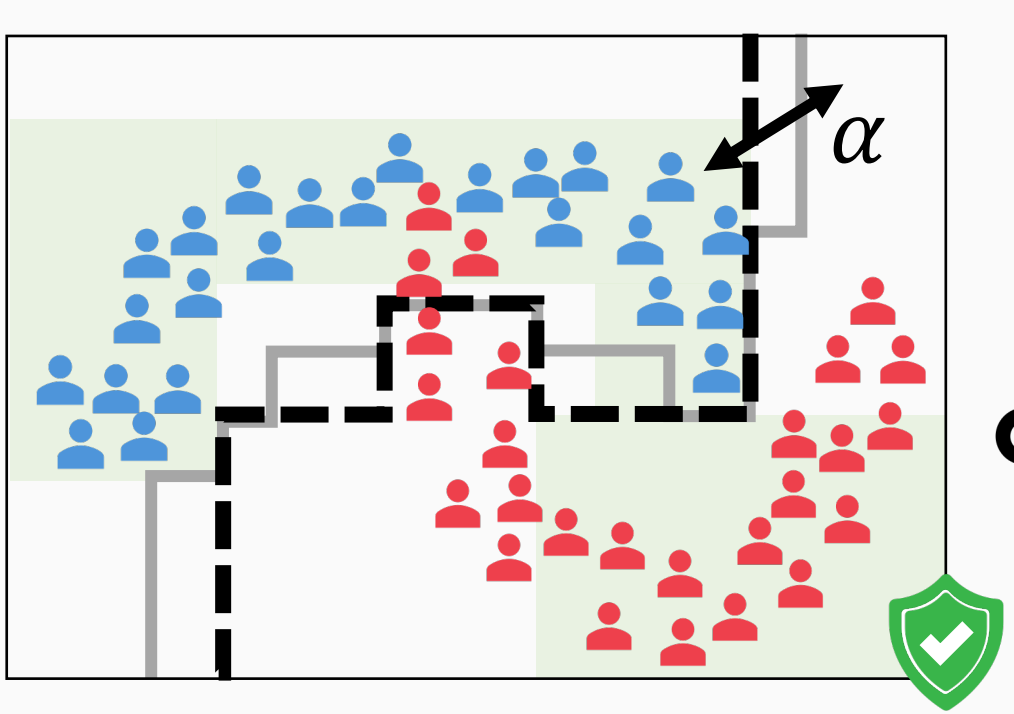
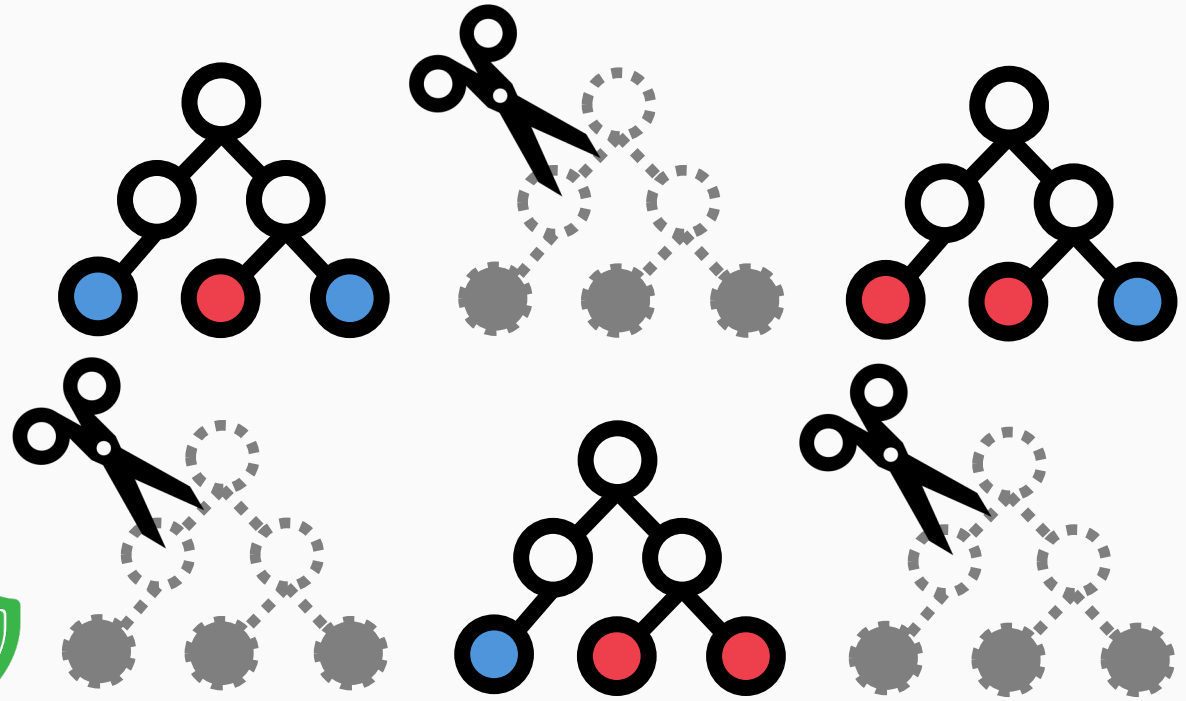
東京大学
THE UNIVERSITY OF TOKYO

ICML
International Conference
On Machine Learning

PINE: Pruning Boosted Tree Ensembles with Conformal In-Distribution Prediction Equivalence



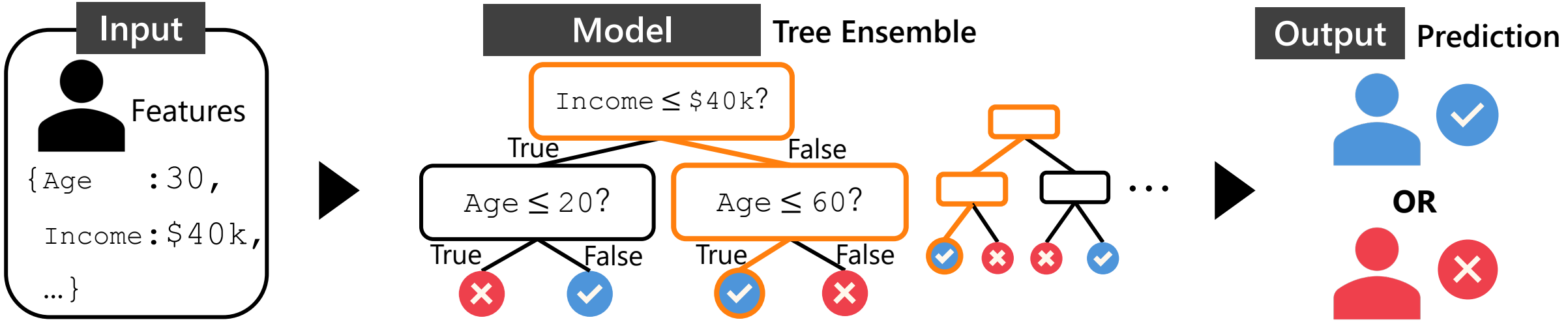


Haruki Yajima, Yusuke Matsui

The University of Tokyo

Problem | Pruning Tree Ensembles

Tree ensemble classification (e.g., XGBoost, LightGBM)



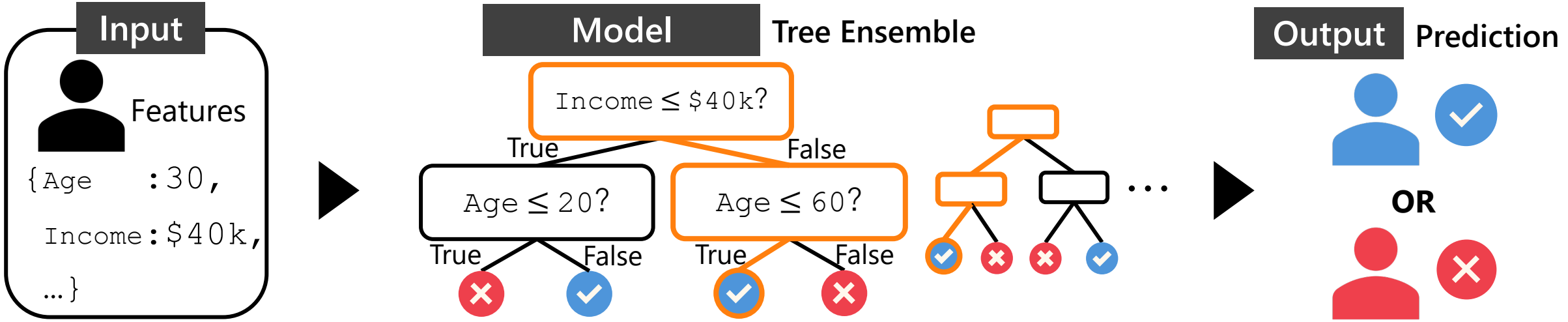
Pruning：Removing Trees

Input Trained Tree Ensemble

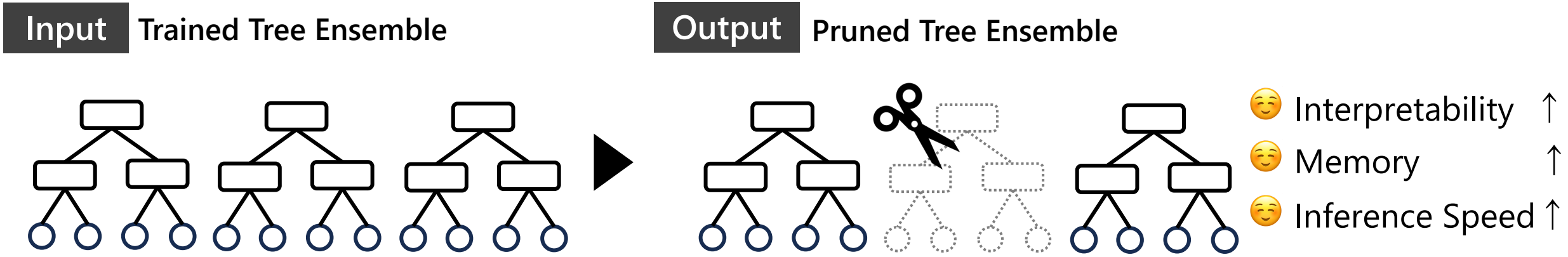
Output Pruned Tree Ensemble

Problem | Pruning Tree Ensembles

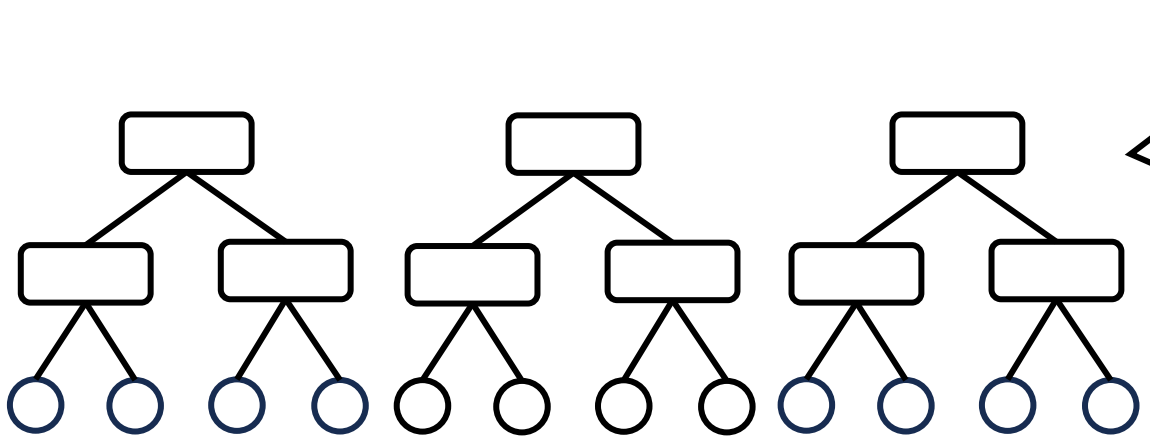
Tree ensemble classification (e.g., : XGBoost, LightGBM)

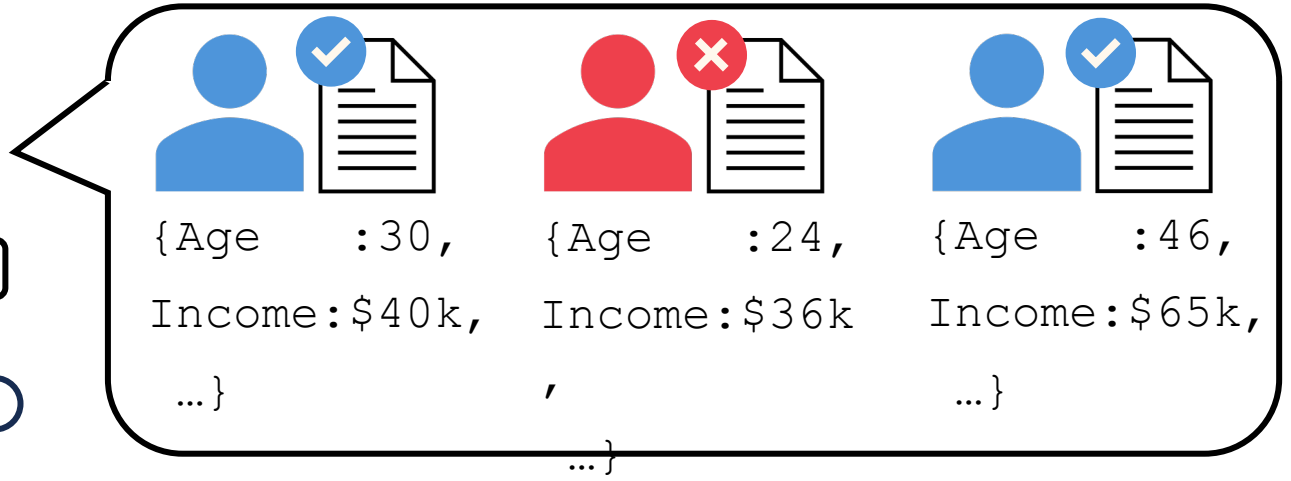


Pruning：Removing Trees

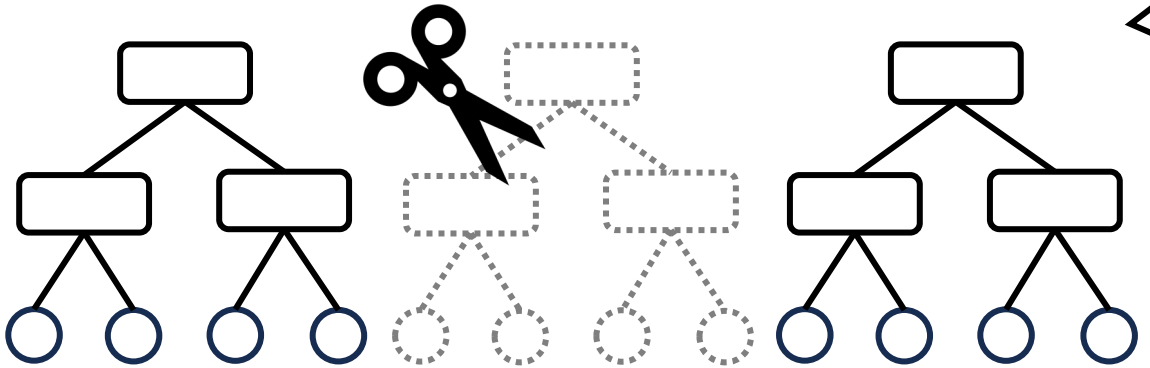


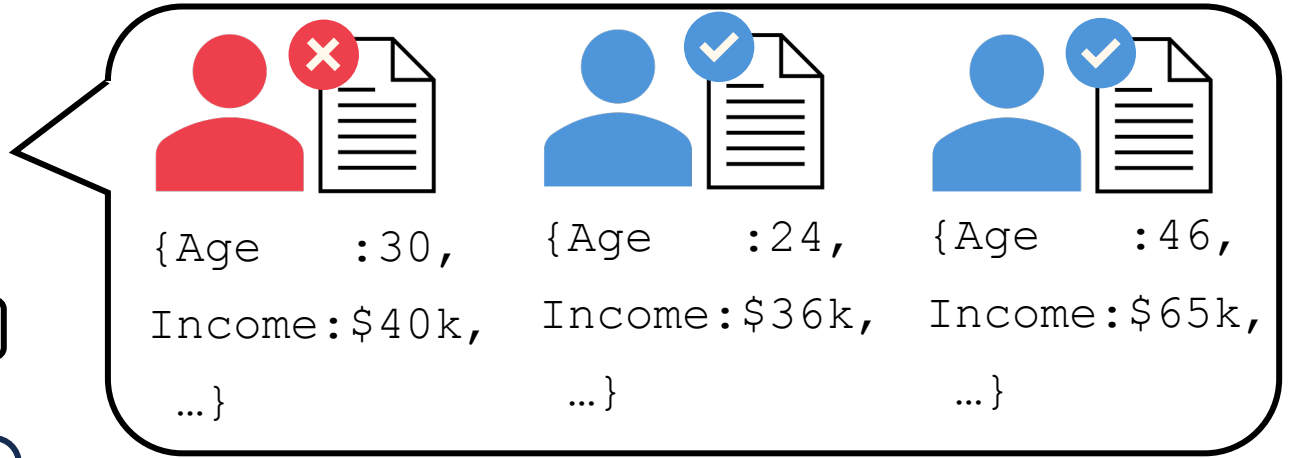
Background | Pruning Can Change Predictions





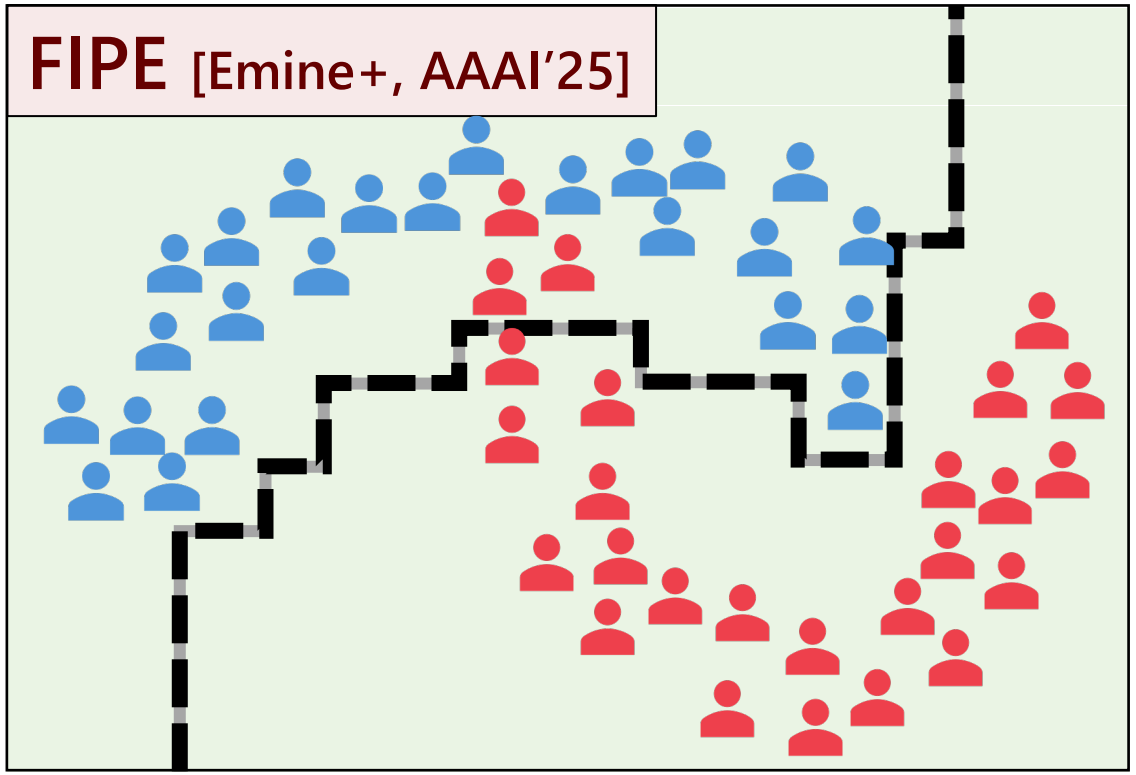
Pruning



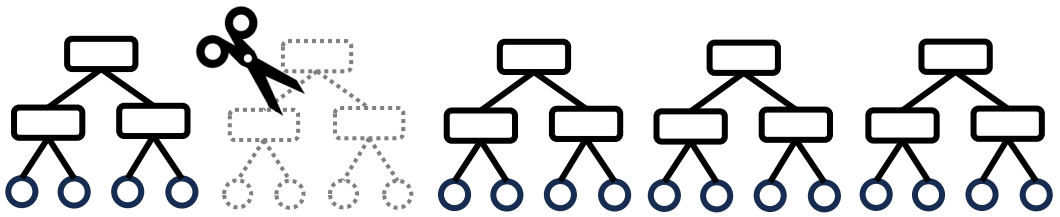


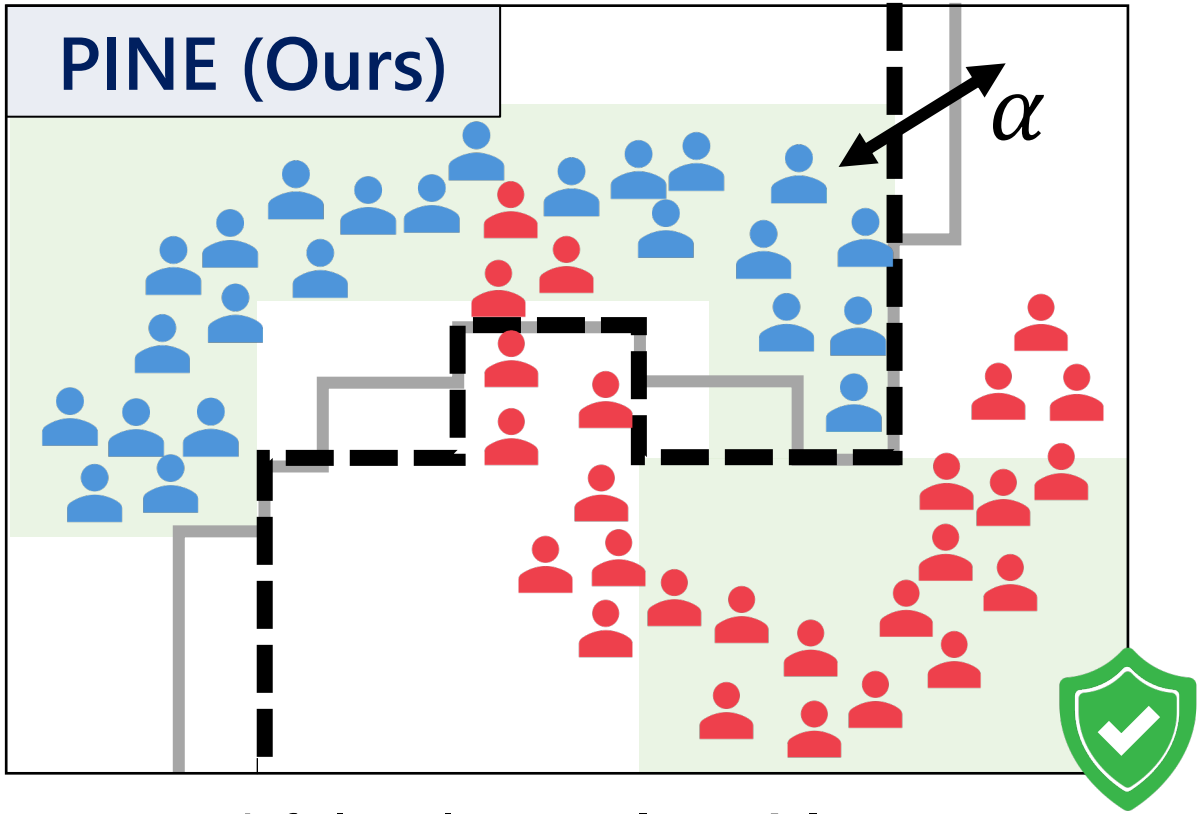
Prediction change

Overview | Faithful Pruning

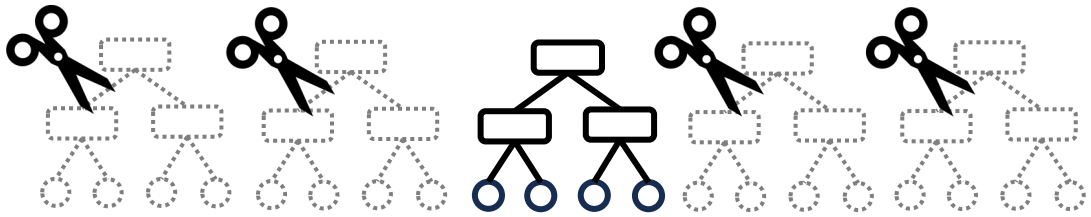


- Faithful on **all inputs**
- **Many** trees retained





- Faithful only on **plausible inputs**
- **Fewer** trees retained

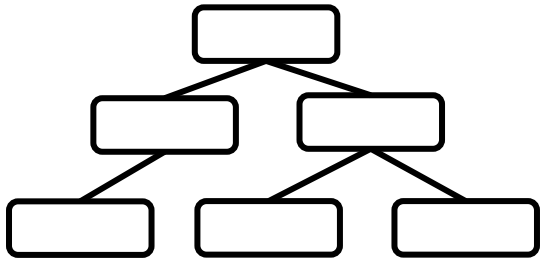


Key idea

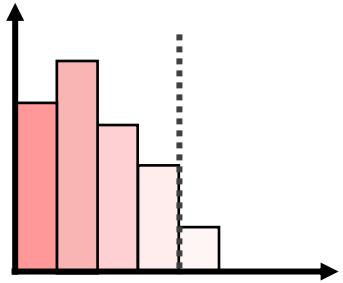
💡 **guarantee faithfulness only on plausible inputs**

$$\underbrace{s(\boldsymbol{x})}_{\text{Plausible score } s(\boldsymbol{x})} \leq \underbrace{\tau(\alpha)}_{\text{Threshold } \tau(\alpha)}$$

Chow-Liu tree



Split Conformal

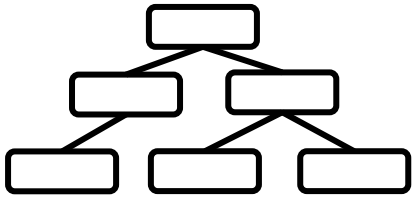


$$\begin{aligned} &\operatorname{argmin} && \#\text{retained trees} \\ &\text{subject to} && \text{Faithful@In-distribution region } \boldsymbol{\mathcal{X}}_{\mathrm{ID}}(\alpha) \end{aligned}$$

Method | Plausible score

Plausible score $s(\boldsymbol{x})$

Chow-Liu tree



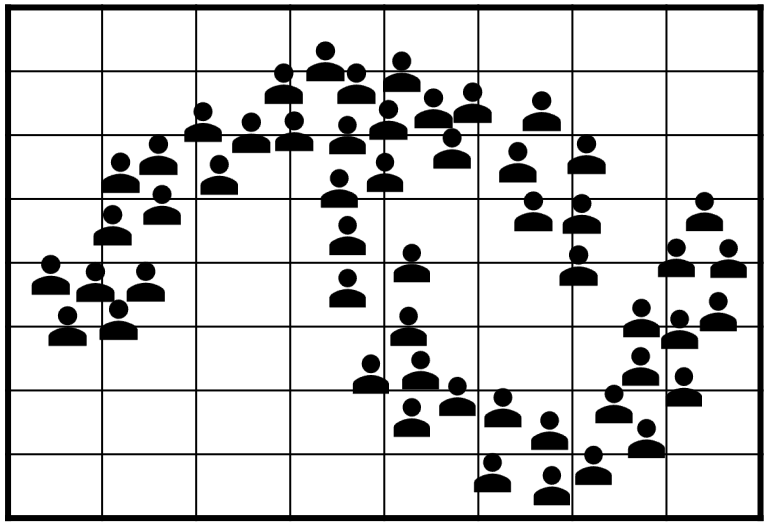
Threshold $\tau(\alpha)$

Split Conformal

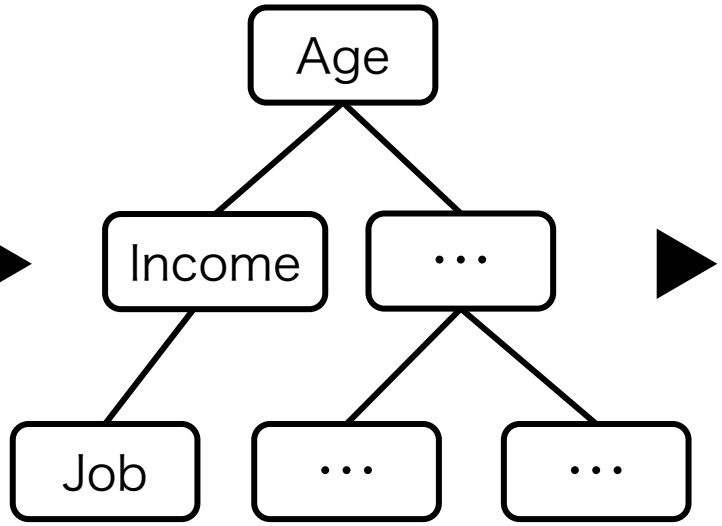
Pruning @$\mathcal{X}_{\mathrm{ID}}(\alpha)$

FIPE

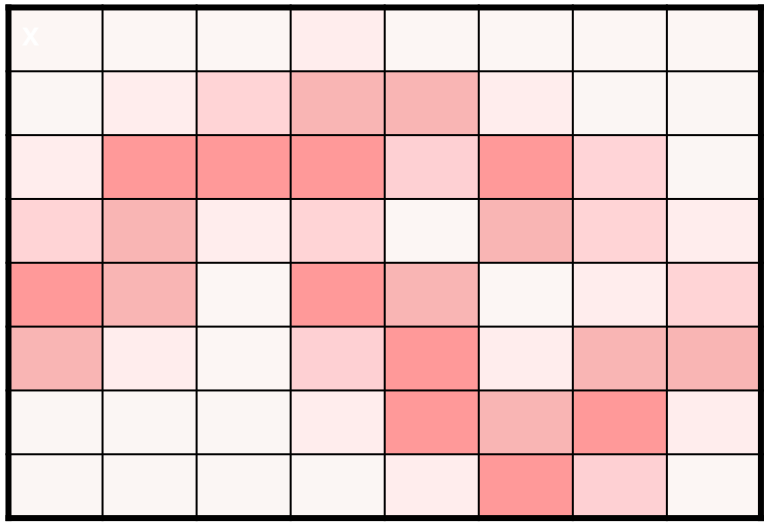
A lightweight distribution model using a Chow-Liu tree [Chow+, IEEE TIT'68]



Input Fit set $\mathcal{D}_{\mathrm{fit}}$



Model Chow-Liu tree



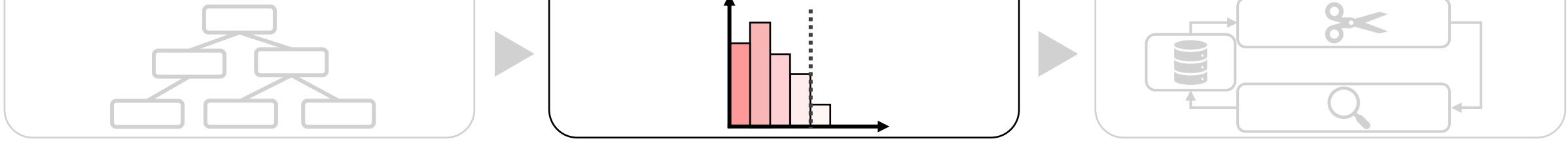
Output Region-wise plausibility

Capture strongest feature dependencies

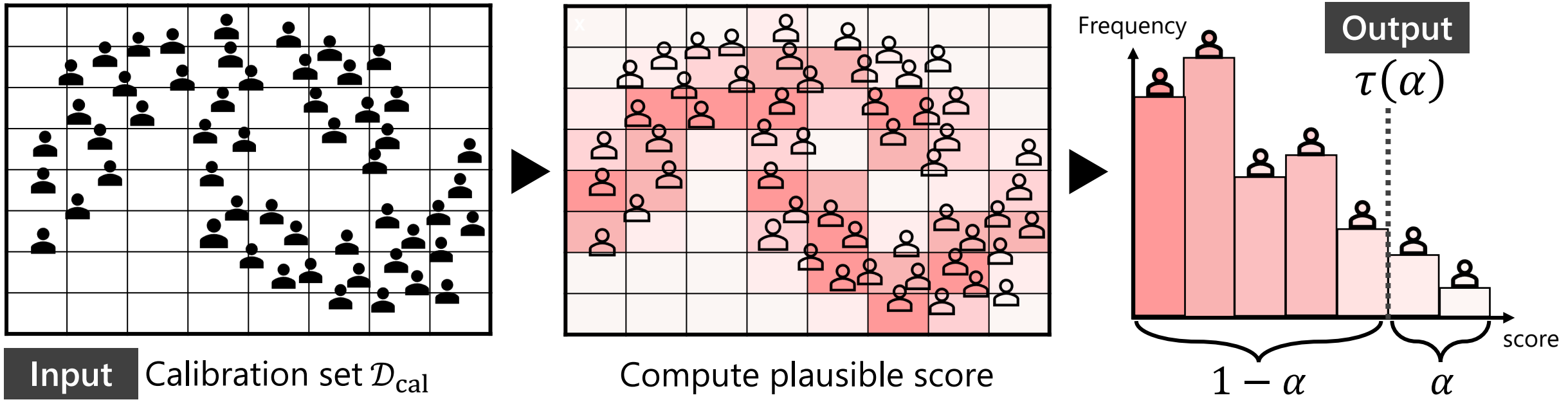
Method | Threshold Calibration

Plausible score $s(x)$ | **Threshold $\tau(\alpha)$** | Pruning @$\mathcal{X}_{\mathrm{ID}}(\alpha)$

Chow-Liu tree | Split Conformal | FIPE



Threshold calibration via split conformal calibration [Lei+, JASA'18]



Input Calibration set $\mathcal{D}_{\mathrm{cal}}$

Compute plausible score

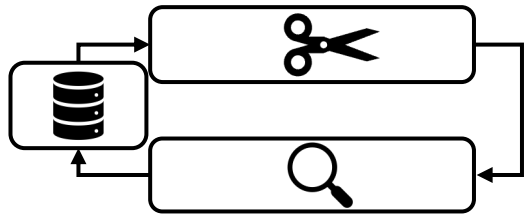
Set $\tau(\alpha)$ as the $(1-\alpha)$-quantile

Method | In-Distribution Faithful Pruning

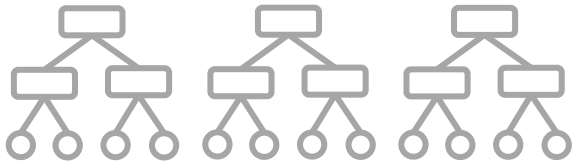
Plausible score $s(x)$ — Chow-Liu tree

Threshold $\tau(\alpha)$ — Split Conformal

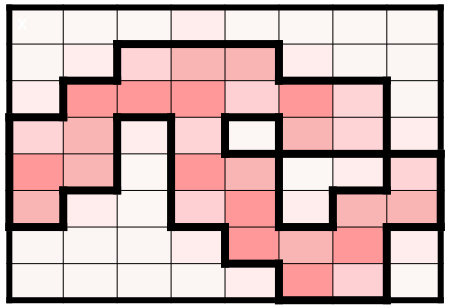
Pruning @$\mathcal{X}_{\mathrm{ID}}(\alpha)$ — FIPE



Apply FIPE [Emine+, AAAI'25] within the in-distribution region



Input Original Ensemble

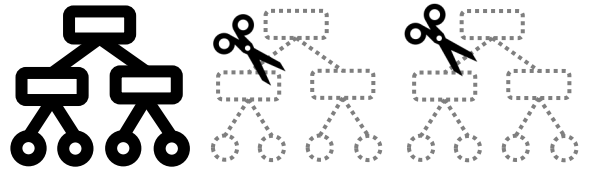


Input In-distribution region $\mathcal{X}_{\mathrm{ID}}(\alpha)$

FIPE-based pruning

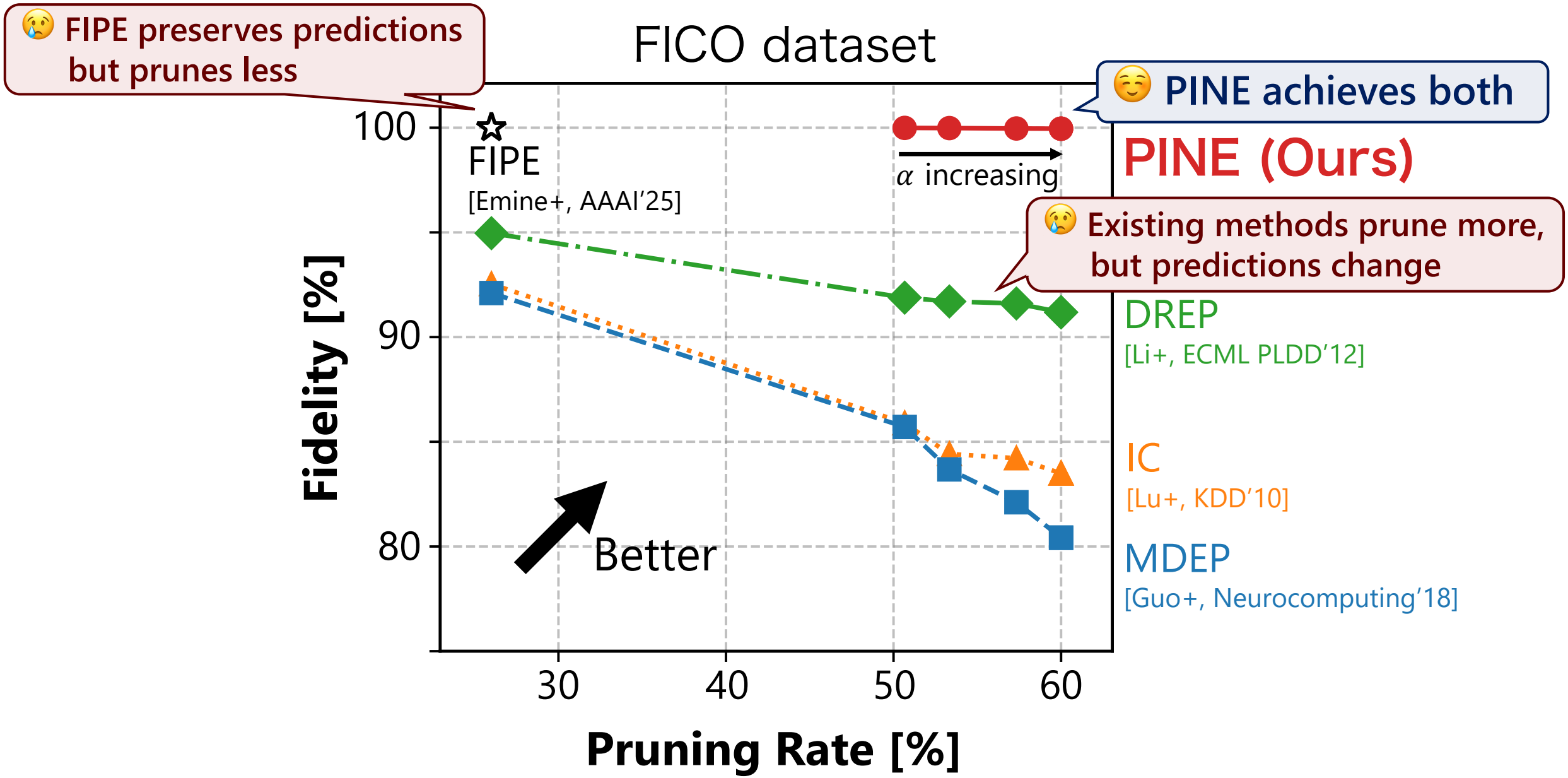
Prune as much as possible

Search for counterexamples
With in plausible region

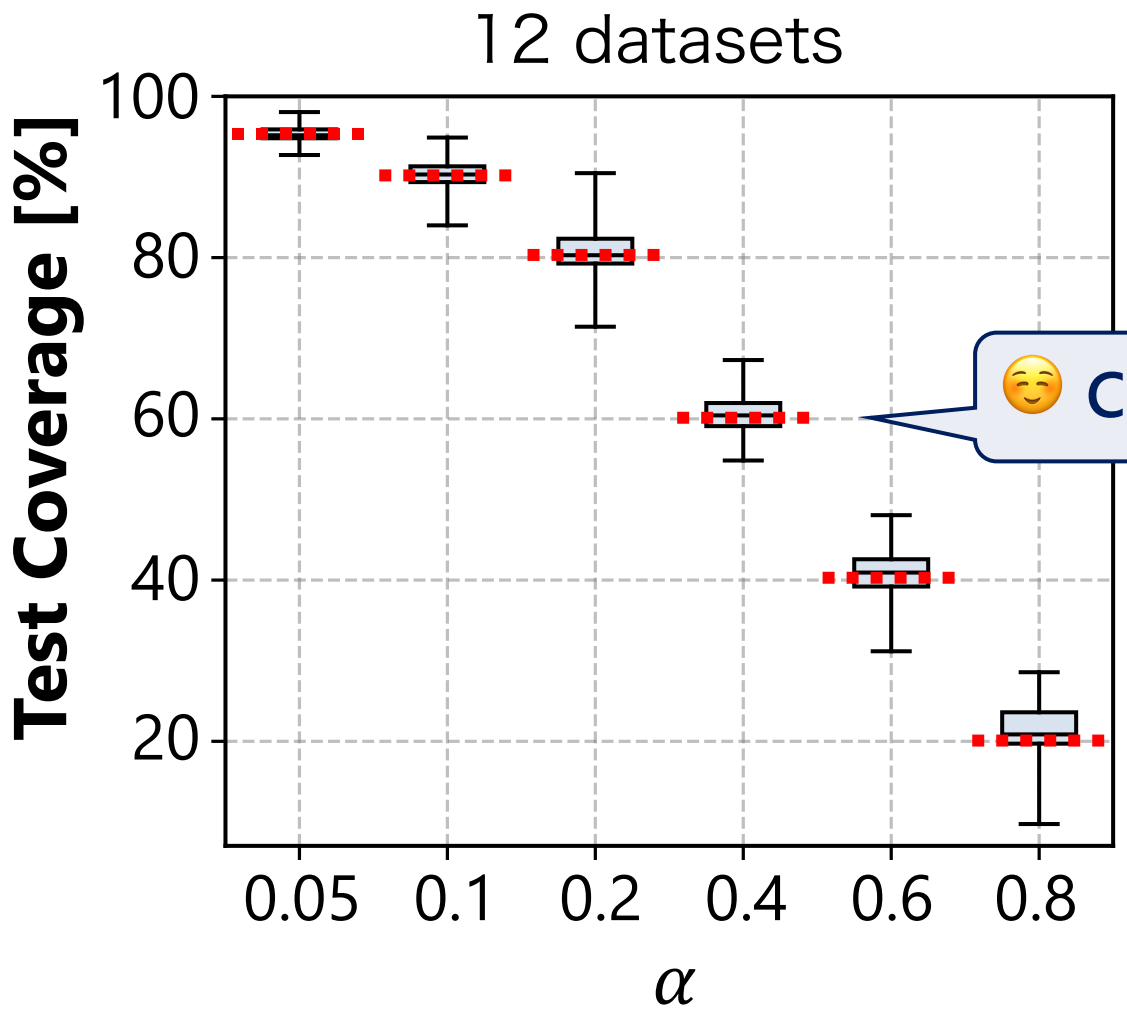


Output Pruned Ensemble

Results | Better Pruning-Fidelity Trade-off



Results | Coverage Control



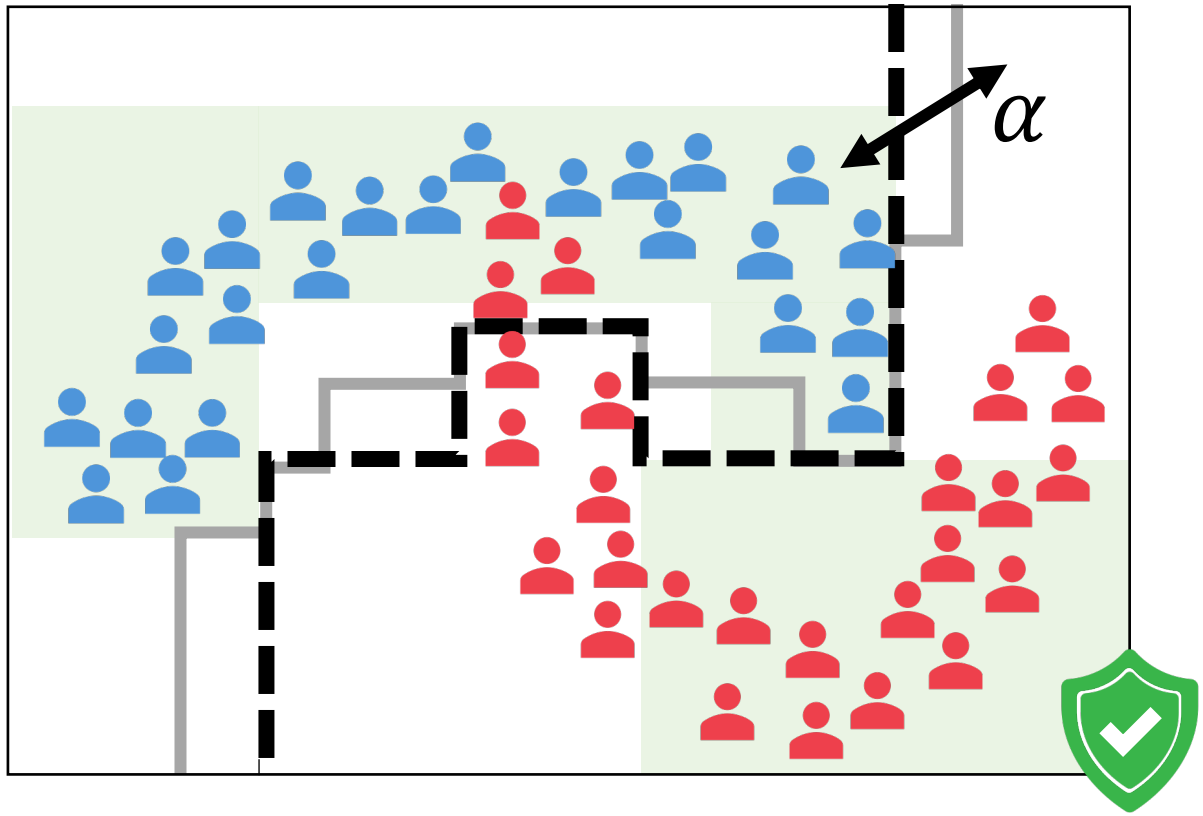
Coverage of the faithful region

: Target Coverage $1 - \alpha$

☺ **Coverage closely matches the target**

Takeaways

- **Pruning tree ensembles while preserving predictions**
  - Improves the pruning-fidelity trade-off across 12 datasets



- ☺️ **Pros**
  - A single parameter controls the faithful region
  - Useful when ground-truth labels are scarce
- 😢 **Cons**
  - No guarantee under distribution shift
  - Predictions may change on OOD inputs