40th Annual Conference on Neural Information Processing Systems (NeurIPS2026)

Distributionally Robust Algorithmic Recourse for Tree Ensembles

Kentaro Kanamori[1], Ken Kobayashi[2], Takuya Takagi[1]

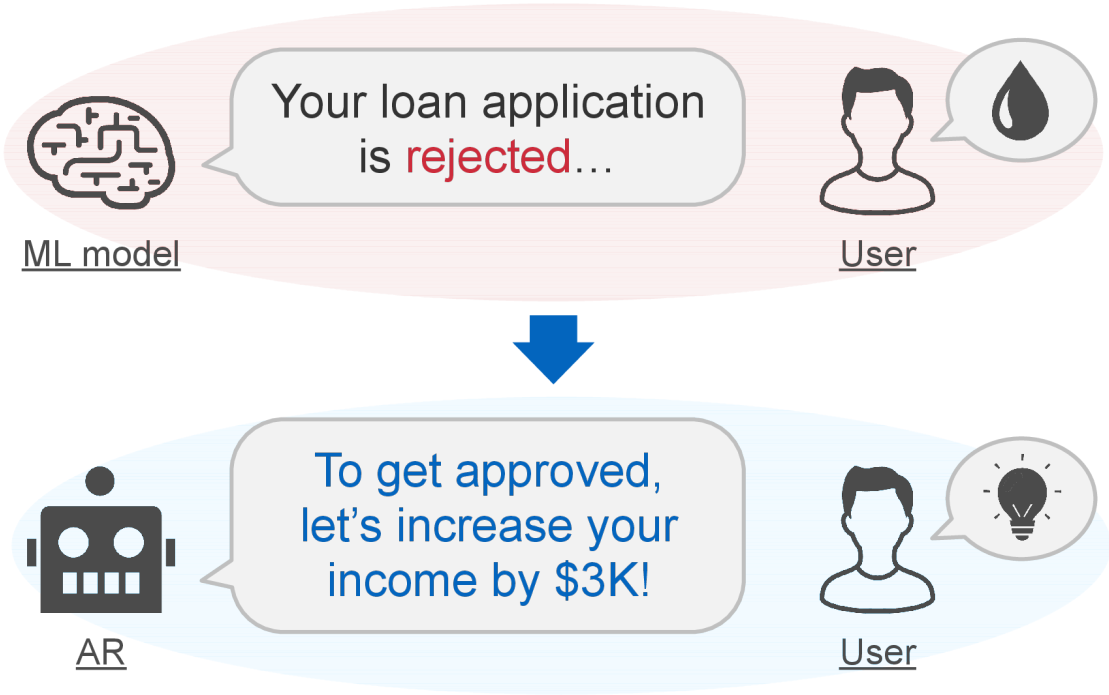
[1]Fujitsu Limited [2]Institute of Science Tokyo

Background

Algorithmic recourse aims to provide "actions" for altering undesired predictions

Algorithmic Recourse [Ustun+ 19]

Explaining a "***recourse action***" for obtaining the desired prediction from an ML model



Algorithmic Recourse (AR) [Ustun+ 19]

Given a model $h: \mathcal{X} \to \mathcal{Y}$, an input instance $x \in \mathcal{X}$, and a desired class $y^* \in \mathcal{Y}$, find an action a^* such that

$$a^* = \arg\min_{a \in \mathcal{A}(x)} c(a \mid x) \quad \text{s.t.} \quad h(x+a) = y^*$$

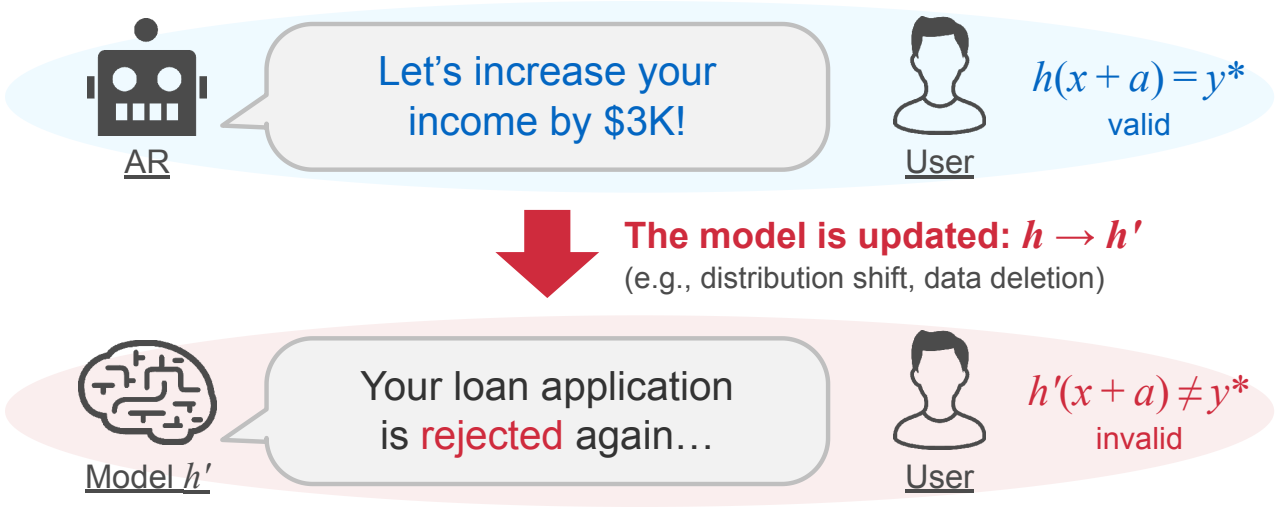
where $\mathcal{A}(x)$ is a set of feasible actions and c is a cost function.

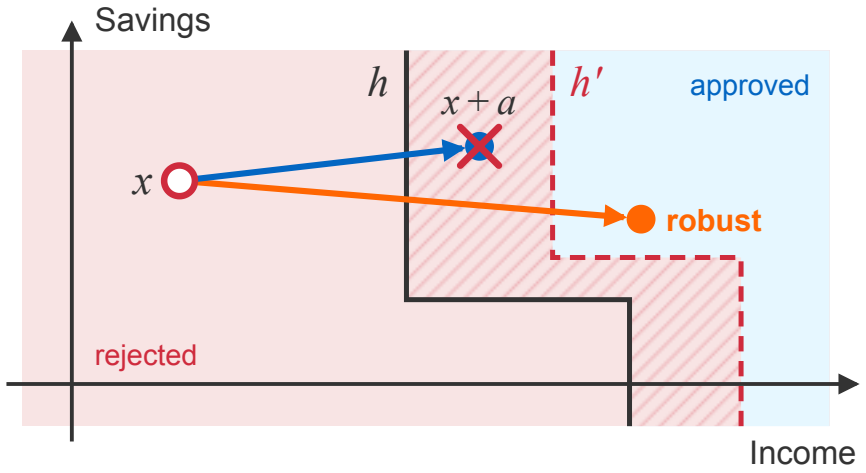
- Provide a low-cost action a that is ***valid*** for the current model h, i.e., $h(x+a) = y^*$

Motivation

Recourse actions may become invalid when the model changes over time

Standard AR methods assume that the model h does not change over time, however…





- Actions optimized for the current model h may not be valid for the updated model h'

Robust AR [Upadhyay+ 21]

Find an action that remains valid for all models in an uncertainty set around the current model h

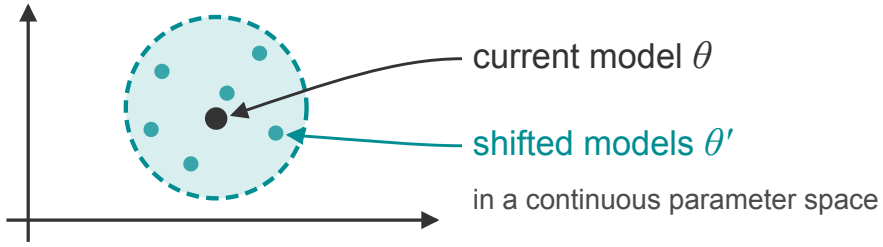
Challenge

Most of existing robust AR methods cannot be directly applied to tree ensembles

We focus on ***tree ensembles***: state-of-the-art models for tabular datasets [Grinsztajn+ 22]

Existing Robust AR

Define the uncertainty set by perturbing continuous model parameters (e.g., weights)

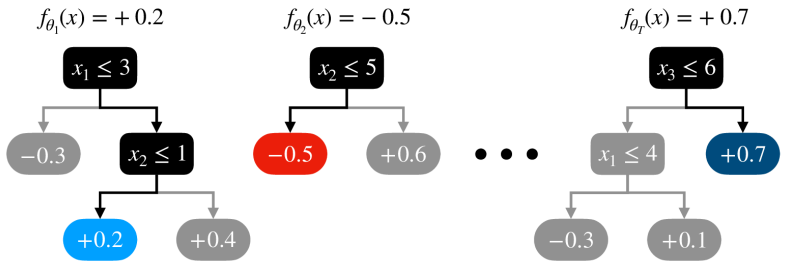


- For linear models and neural networks

Tree Ensembles (e.g., XGBoost, random forests)

Parameters include discrete tree structures

(and thus the model is neither differentiable nor smooth...)



- Cannot be perturbed in the same way

Our Goal

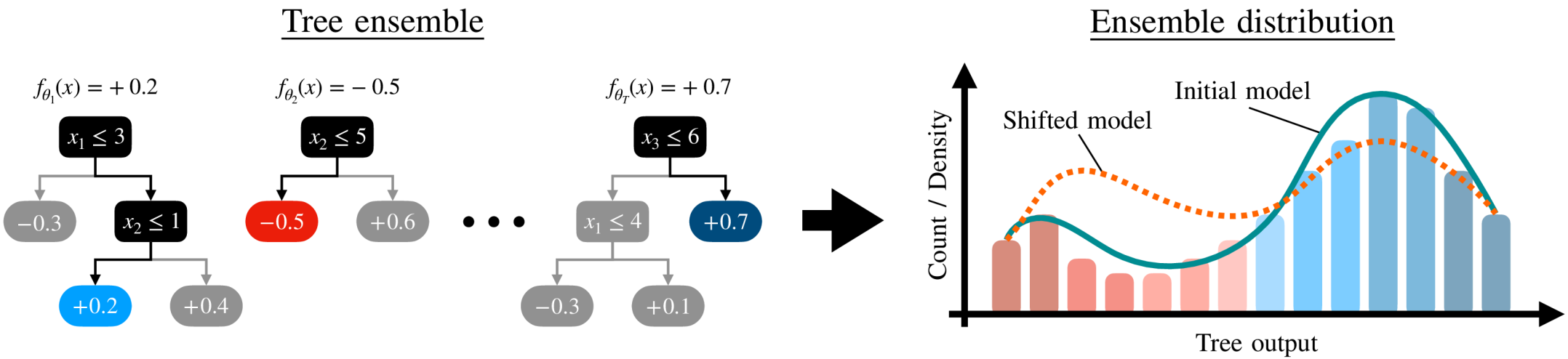
Propose a new framework for providing robust recourse actions for tree ensembles

Our Approach: Overview

Express the model change as a perturbation to the distribution of tree outputs

Ensemble Distribution [Malinin+ 20]

Distribution $q_\theta(\hat{y} \mid x)$ of the tree outputs $f_{\theta_1}(x), \ldots, f_{\theta_T}(x) \implies f_\theta(x) = \mathbb{E}_{\hat{y} \sim q_\theta(\cdot|x)}[\hat{y}]$



Our Idea

Perturb the ensemble distribution q_θ at $x + a$, instead of the model parameter θ itself

Our Approach: Problem Formulation

Formulate robust AR for tree ensembles as distributionally robust optimization

Distributionally Robust AR for Tree Ensembles (DiRARTE)

Given an instance x, a tree ensemble f_θ, and a radius $\varepsilon \geq 0$, we consider the following problem:

$$\min_{a \in \mathcal{A}(x)} \; c(a \mid x) \quad \text{s.t.} \quad \min_{q \in \mathcal{Q}_\varepsilon(\theta \mid x+a)} \mathbb{E}_{\hat{y} \sim q} \left[\rho(y^*, \hat{y}) \right] \geq 0$$

Valid for all the distributions in the uncertainty set

where $\mathcal{Q}_\varepsilon(\theta \mid x + a) := \{q \mid \Delta(q, q_\theta(\cdot \mid x + a)) \leq \varepsilon\}$ is the *uncertainty set* at $x + a$ with a discrepancy measure Δ, and $\rho(y^*, \hat{y}) := y^* \cdot \hat{y}$ is the margin.

- Can express robustness to model changes without explicitly constructing shifted models
- However, the constraint involves a minimization over distributions (generally intractable)

Our Approach: Tractable Reformulation

Our robust validity constraint can be evaluated as efficiently as the standard one

Wasserstein Distance

Worst case: shift all the tree outputs



Proposition 1. (Wasserstein distance of order $p \geq 1$)

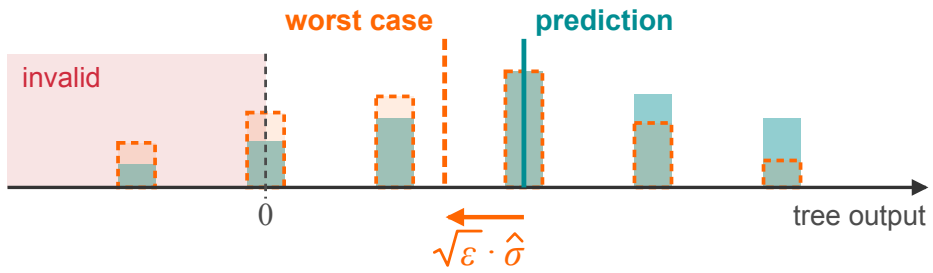
Our constraint with the Wasserstein distance is equivalent to

$$\rho(y^*, f_\theta(x+a)) \geq \varepsilon$$

- Require the margin to be at least the radius ε

χ^2-Divergence

Worst case: reweight the tree outputs



Proposition 2. (χ^2-divergence; for sufficiently small ε)

Our constraint with the χ^2-divergence is equivalent to

$$\rho(y^*, f_\theta(x+a)) \geq \sqrt{\varepsilon} \cdot \hat{\sigma}_\theta(x+a)$$

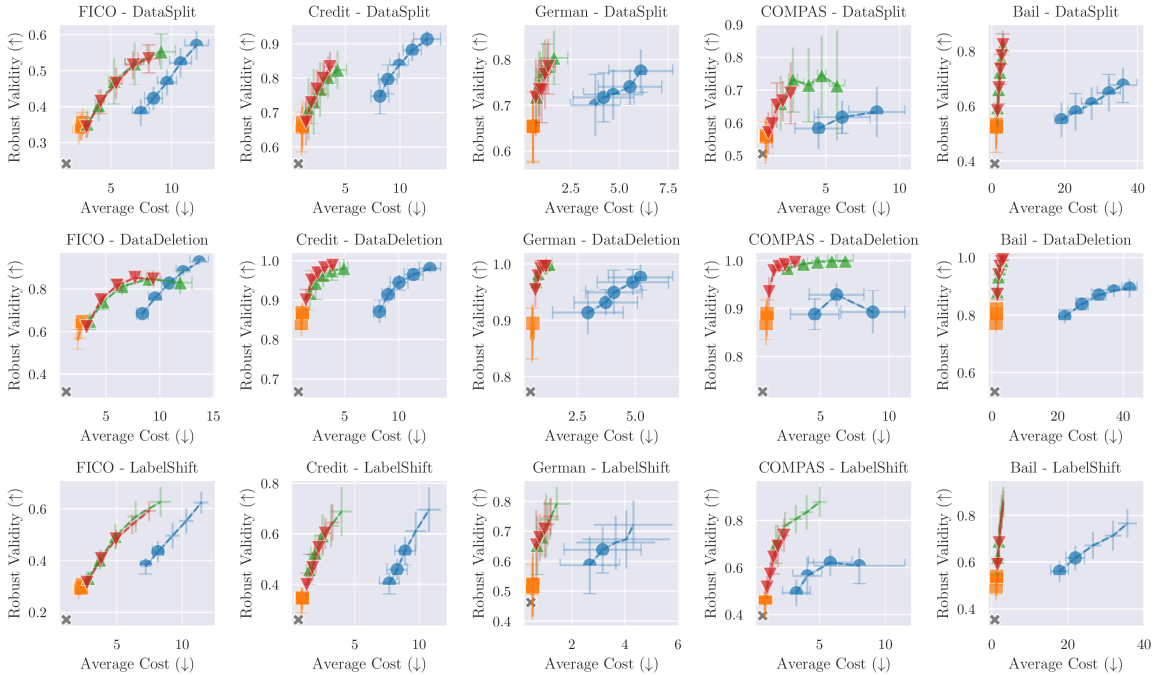
- Require a larger margin for more diverse trees

- Can be efficiently optimized by our FT++, an extension of the feature tweaking algorithm [Tolomei+ 17]

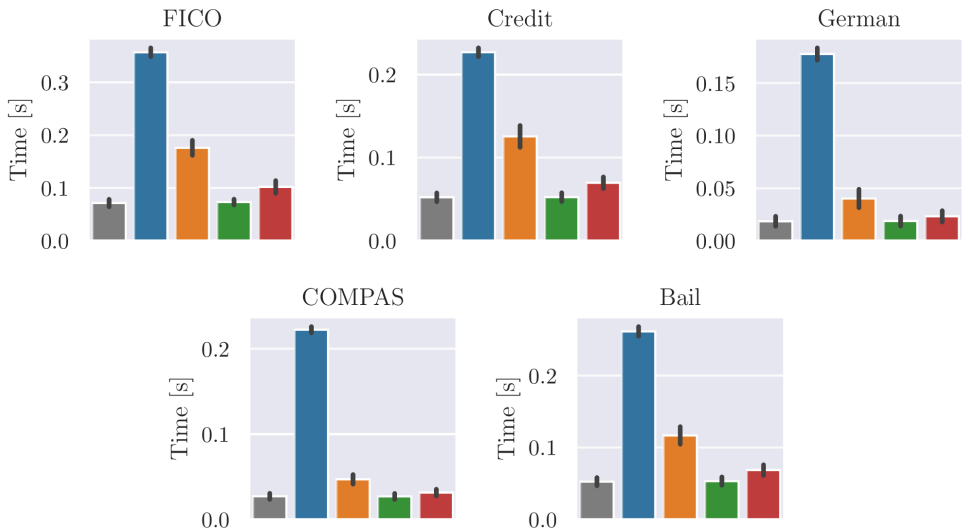
Experiments

Achieved a better robustness–cost trade-off while maintaining comparable runtime

Exp. 1 Robustness–Cost Trade-off (upper left is better)



Exp. 2 Runtime per Instance (lower is better)



Baseline (non-robust) | RobX [Dutta+ 22] | BetaRobust [Stępka+ 25] | DiRARTE-W (ours) | DiRARTE-χ² (ours)

- ***DiRARTE***: lower cost than RobX in most cases; runtime comparable to the non-robust baseline

Summary

Providing recourse actions robust to model changes in tree ensembles

We propose ***Distributionally Robust AR for Tree Ensembles (DiRARTE)***:

- Formulate robust AR for tree ensembles as a DRO problem
  - Express model changes by perturbing the ensemble distribution
- Derive tractable reformulations of the robust validity constraint
  - Can be evaluated as efficiently as the standard validity constraint
- Demonstrate the efficacy of our method by experiments on real datasets
  - Achieve a better robustness–cost trade-off with comparable runtime



OpenReview



GitHub