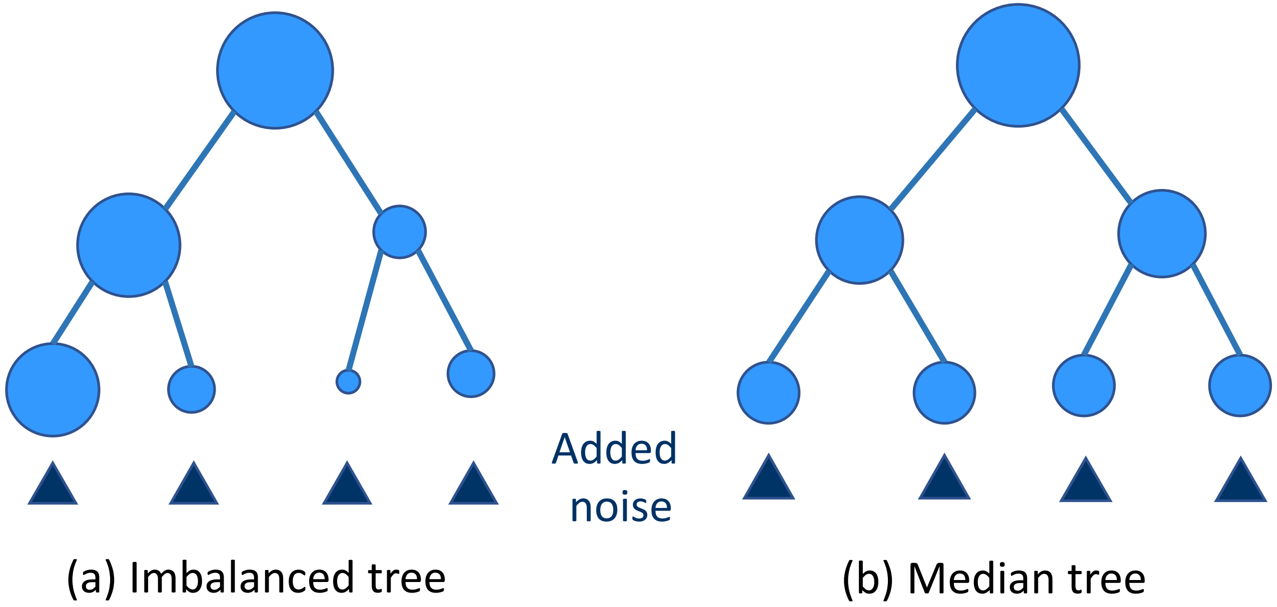
# Differentially Private Random Forests for Regression and Classification

Shorya Consul and Sinead A. Williamson, UT Austin

- Adding privacy-protecting noise to low-occupancy leaf nodes overwhelms the signal.

- DiPriMe is based on the idea of splitting on private medians → learn more balanced trees

- We derive theoretical utility bounds and demonstrate state-of-the-art performance on multiple datasets



(a) Imbalanced tree

(b) Median tree

Existing tree-based methods learn trees like (a) while DiPriMe leads to trees like (b). The noise added during privatization swamps the “signal” at low-occupancy leaves in (a). Median splits resolves this by learning more balanced leaves.

TEXAS
The University of Texas at Austin