

Topic 20: Random Forest

by Sai Zhang

Key points: .

Disclaimer: *The note is built on Prof. [Jinchi Lv](#)'s lectures of the course at USC, DSO 607, High-Dimensional Statistics and Big Data Problems.*

20.1 Motivation

Denote by $m(\mathbf{X})$ the measurable nonparametric regression function with p -dimensional random vector \mathbf{X} taking values in $[0, 1]^p$. The Random Forest algorithm aims to learn the regression function in a non-parametric way based on the observations $\mathbf{x}_i \in [0, 1]^p, y_i \in \mathbb{R}$

[Chi et al. \(2022\)](#)

References

Chien-Ming Chi, Patrick Vossler, Yingying Fan, and Jinchi Lv. Asymptotic properties of high-dimensional random forests. *The Annals of Statistics*, 50(6):3415–3438, 2022.