Decision Tree

Shiqiang Jin

September 27, 2019

Consider a continuous response variable $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ and p many predictors $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p) \in \mathbb{R}^{n \times p}$.

Algorithm 1 Rough algorithm of decision tree

---

1. For each predictor $\mathbf{x}_j$, we partition it into two distinct regions $R_1(j, s) = \{\mathbf{x}_j | \mathbf{x}_j < s\}$ and $R_2(j, s) = \{\mathbf{x}_j | \mathbf{x}_j \geq s\}$.

2. For each predictor $\mathbf{x}_j$, observations are divided into two regions R_1 or R_2, then we make the same prediction with $\hat{\mathbf{y}}_{R_1} = \frac{1}{n_1} \sum_{i \in R_1(j,s)} y_i$ or $\hat{\mathbf{y}}_{R_2(j,s)} = \frac{1}{n_2} \sum_{i \in R_2} y_i$.