

清华大学

TSINGHUA UNIVERSITY

① ID3: 选择最大信息增益 (Gini index) (熵)
 $IG = H(D) - \sum \frac{|D_i|}{|D|} H(D_i)$
 Greedy 选择 (不能 do it)

② 决策树: 选择最大信息增益 (Gini index) (熵)
 $H(x) = -\sum p(x_i) \log p(x_i)$
 $P(H(x)=y) = P(\sum I(h_i(x), y) \geq \frac{1}{2})$
 $P(H(x)=y) = P(\sum I(h_i(x), y) \geq \frac{1}{2})$
 $\leq 2 \exp(-\frac{1}{2} \sum (I(h_i(x), y) - \frac{1}{2})^2)$

③ CART: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

Random forest: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

④ 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑤ CART: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑥ Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑦ Gradient Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑧ Gradient Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑨ Gradient Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑩ Gradient Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑪ Gradient Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)

⑫ Gradient Boosting: 选择最大信息增益 (Gini index) (熵)
 1) 选择最大信息增益 (Gini index) (熵)
 2) 选择最大信息增益 (Gini index) (熵)
 3) 选择最大信息增益 (Gini index) (熵)



