28/08/25 (AIL7024)

Boardwork (DTS)

longer DT

Training | validation | Test
(20%)

1. during Training → Specify a depth
   → IG threshold
   → Statistical Significance test (Chi sq., t-test)
   → Minimum Description Length (MDL)
2. post Training
   ↓
   pruning

—— ∝ ——

Rule-post pruning

(cond^n 1) ∧ (cond^n 2) ∧ (cond^n 3) → Y

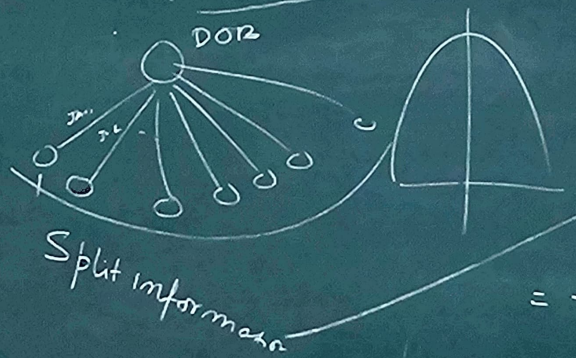
context

—— ∝ ——

		t_1			t_2	
Temp:	40	48	60	72	80	90
Playtime:	N	N	Y	Y	Y	N

IG(temp t_1) IG(temp t_2)

DOB Overcome Overfitting Problem Training / Validation / DRV

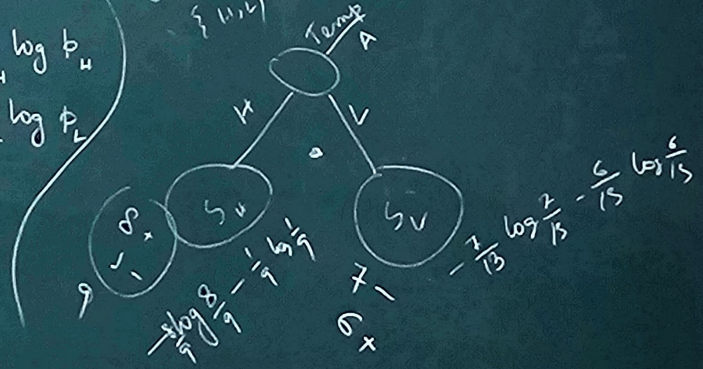


$$IG(S,A) = H(S) - \sum_{v \in Value(A)} H(S_v) * \frac{|S_v|}{|S|}$$

$$SI(A)$$

$$H(S,A) = -p_H \log p_H - p_L \log p_L$$

$$= -\frac{9}{22}\log\frac{9}{22} - \frac{13}{22}\log\frac{13}{22}$$



DOB Overcome overfitting Problem

probability.

Dia — NO

(6) Temp - Hot — 3/6
Mild — 2/6
Cold — 1/6

Temp
Dia — $\left(\frac{1}{2}H, \frac{1}{3}M, \frac{1}{6}C\right)$