

AI-Assignment -7

19K41A04H6

Given data :-

outlook	Temperature	Humidity	windy	Houmpby
Rainy	Hot	High	False	25
Rainy	Hot	High	True	30
overcast	Hot	High	False	46
Sunny	Mild	High	False	45
Sunny	Cool	Normal	False	32
overcast	Cool	Normal	True	23
Rainy	Cool	Normal	True	43
Rainy	Mild	Normal	False	35
Sunny	Cool	High	False	38
Rainy	Mild	Normal	False	46
overcast	Mild	Normal	False	48
Overcast	Hot	Normal	True	52
Sunny	Mild	Normal	False	44
Decision tree for the given data.				30

Target table (PGH) :-

$$\text{mean} = 39.78, \text{SD}(T) = 9.32 \quad \therefore CV = \frac{\text{SD}}{\text{mean}} \times 100 = 23$$

Outlook

→ Rainy → 8.7	→ 5	weight
→ overcast → 4.03	4	5/14
→ sunny → 12.15	5	4/14

$$SD(\text{outlook}) = \left[\frac{5}{14} \times 8.7 \right] + \left[\frac{4}{14} \times 4.03 \right] + \left[\frac{5}{14} \times 12.15 \right]$$

$$= 8.59$$

$$SDR(\text{outlook}) = \text{SD}(T) - \text{SD}(\text{outlook})$$

$$= 9.32 - 8.59 = 0.73$$

Temperature

→ Hot → 10.3410	4	weight(x)
→ Mild → 8.38	6	4/14
→ cool → 12.13	4	6/14
		4/14

$$SD(T \text{ temperature}) = \left\{ \frac{4}{14} * (0.34) \right\} + \left\{ \frac{6}{14} * 8.32 \right\} + \left\{ \frac{4}{14} * \text{target} \right\}$$

$$= 10.01$$

$$SDR(\text{temperature}) = SD(T) - SD(\text{temperature})$$

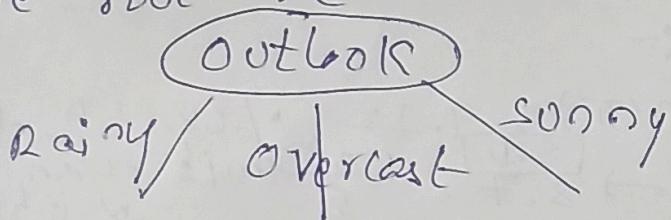
$$= 9.32 - 10.01 = -0.78$$

humidity → $\begin{cases} \text{High} \rightarrow 9.5169 - 7 & \frac{7}{14} \\ \text{Normal} \rightarrow 9.433 - 7 & \frac{7}{14} \end{cases}$

$$SD(\text{humidity}) = \frac{7}{14} * 9.51 + \frac{7}{14} * 9.43 = 9.47$$

$$\begin{aligned} SDR(\text{Humidity}) &= SD(T) - SD(\text{humidity}) = 9.32 - 9.47 \\ &= -0.15 \end{aligned}$$

→ The Outlook has the highest SDR so it will become root node



Rainy table:

Temp	Humidity	Windy	Hours to play
Hot	High	False	25
Hot	High	True	30
Mild	High	False	35
Cool	Normal	False	30
Mild	Normal	True	48

target column (PHT)

$$\text{mean} = 35.0 \quad \text{SD} = \frac{\sum}{\text{mean}} \times 100 = 24.7 \\ \text{SD(T)} = 8.70$$

weight

2/5

Temp	Hot	$\rightarrow 1.53 \rightarrow 2$	2/5
	mild	$\rightarrow 9.19 \rightarrow 2$	
	cool	$\rightarrow 0 \rightarrow 1$	

$$\text{SD(Temp)} = \left[\frac{2}{5} * 1.53 \right] + \left[\frac{2}{5} * 9.19 \right] + \left[\frac{1}{5} * 0 \right] \\ = 5.37$$

$$\text{SOR(Temp)} = \text{SD(Temp)} - \text{SD(Target)} \\ = 5.37 - 8.70 = 3.33$$

Humidity	High	$\rightarrow 7.07 / 3$	3/5
	Normal	$\rightarrow 7.07 / 2$	

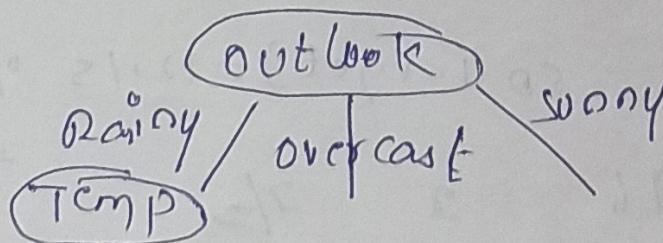
$$\text{SD(Humidity)} = \left[\frac{3}{5} * 7.07 \right] + \left[\frac{2}{5} * 7.07 \right] = 7.07$$

$$\text{SOR} = \text{SD(T)} - \text{SD(Hum)} \\ = 8.70 - 7.07 = 1.63$$

Windy	false	$\rightarrow 9.19 / 3$	3/5
	true	$\rightarrow 12.72 / 2$	

$$\text{SD(Windy)} = \left[\frac{3}{5} * 9.19 \right] + \left[\frac{2}{5} * 12.72 \right]$$

$$\text{SOR} = \text{SD(Target)} - \text{SD(Windy)} \\ = 8.70 - 11.03 = -2.33$$



Overcast table:-

Temperature	Humidity	Windy	Hours to play
Hot	High	False	46
Cool	Normal	True	43
Mild	High	True	52
Hot	Normal	False	44

Target hours to play :-

$$SQT = 46/5 + \dots = 4.0 \geq 11, \text{ mean} = 46.25$$

$$\cdot ICL(T) = \frac{SD}{\text{mean}} \times 100 = 8.71$$

Here the $\cdot ICL(T) \leq 10$, number of samples = 4
so there it contains leaf node

Sunny table:-

Temp	Hum	Windy	Hours to play
Mild	High	False	45
Cool	High	False	52
Cool	Normal	True	23
Mild	Normal	False	46
Mild	High	False	20

Target column :-

$$\text{Mean} = 29.2 \quad SQT(\text{target}) = 12.15 \quad \%CV = \frac{SD}{\text{mean}} \times 100 = 20.9$$

$$\text{Temp} \left[\begin{array}{l} \text{Mild} \rightarrow 8.96 \\ \text{Cool} \rightarrow 20.50 \end{array} \right] \quad \begin{matrix} 3 & 7/5 \\ 2 & 2/5 \end{matrix}$$

$$SD = 12.57$$

$$SDR = 12.15 - 12.57 = -1.42$$

humidity -

high	- 11.23	3	3/5
normal	- 16.26	2	2/5

$$SD(T) = 13.24$$

$$SDR = 12.15 - 13.24 = -1.09$$

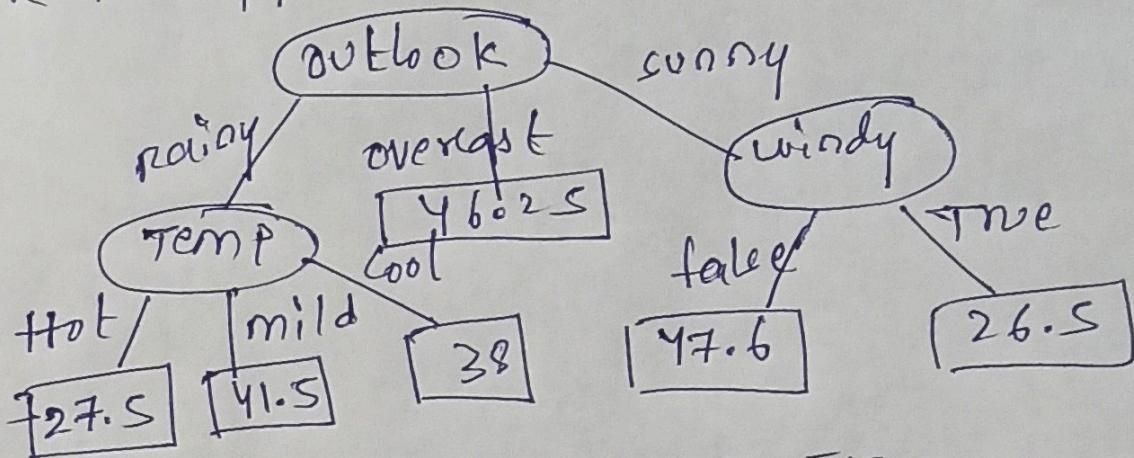
windy -

false	- 3.78	2	3/5
True	- 4.94	2	2/5

$$SD(W) = 4.244$$

$$SDR = SD(T) - SD(W) = 12.15 - 4.244$$

$$SDR = 7.91$$



this is the final True

→ In this problem there is no effect of humidity on the hours of playing