AI ASSIGNMENT - 7

B. Vishnu Priya
19K41A05A1

Given data

Outlook	Temperature	Humidity	Windy	Hrs to play
Rainy	Hot	High	False	25
Rainy	Hot	High	True	30
Overcast	Hot	High	False	46
Sunny	Mild	High	False	45
Sunny	Cool	Normal	False	52
Sunny	Cool	Normal	True	23
Overcast	Cool	Normal	True	43
Rainy	Mild	High	False	35
Rainy	Cool	Normal	False	38
Sunny	Mild	Normal	False	46
Rainy	Mild	Normal	True	48
Overcast	Mild	High	True	52
Overcast	Hot	Normal	False	44
Sunny	Mild	High	True	30

Decision tree for the given data:

Target table (PGH):-

mean = 39.78

SD(T) = 9.32

% CV = $\frac{SD}{mean} \times 100 = 23\%$

5F1 SAI 10

L8111121

		S.D	No. of samples	Weight
Outlook	Rainy	8.7	5	5/14
	Overcast	4.03	4	4/14
	Sunny	12.15	5	5/14

$$S.D(\text{outlook}) = \left[\frac{5}{14} \times 8.7\right] + \left[\frac{4}{14} \times 4.03\right] + \left[\frac{5}{14} \times 12.15\right]$$
$$= 8.59$$

$$SDR(\text{outlook}) = S.D(T) - S.D(\text{outlook})$$
$$= 9.32 - 8.59$$
$$= 0.73$$

Temperature	Hot	10.340	4	4/14
	Mild	8.38	6	6/14
	Cool	12.13	4	4/14

$$S.D(\text{temperature}) = \left[\frac{4}{14} * 10.34\right] + \left[\frac{6}{14} * 8.38\right] + \left[\frac{4}{14} * 12.13\right]$$
$$= 10.01$$

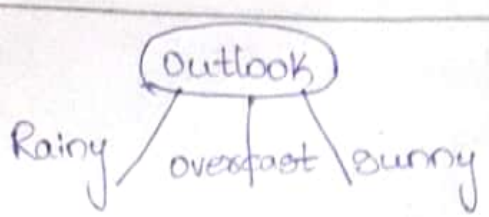
$$SDR(\text{temperature}) = S.D(T) - SD(\text{temperature})$$
$$= 9.32 - 10.01$$
$$= -0.78$$

Humidity	High	9.5169	7	7/14
	Normal	9.433	7	7/14

$$S.D(\text{Humidity}) = \frac{7}{14} * 9.5169 + \frac{7}{14} * 9.433$$
$$= 9.47$$

$$SDR(\text{Humidity}) = S.D(T) - S.D(\text{Humidity})$$
$$= 9.32 - 9.47$$
$$= -0.15$$

* Outlook has the highest SDR. So it will become as root node.



Rainy table :-

Temperature	Humidity	Windy	Hours to play
Hot	High	False	25
Hot	High	True	30
Mild	High	False	35
Cool	Normal	False	38
Mild	Normal	True	48

Target column (PGH) :-

mean = 35.2

% C.V = $\frac{SD}{mean} \times 100 = 24.7$

S.D(T) = 8.70

Temperature
- Hot — 3.53 — 2 — 2/5
- mild — 9.19 — 2 — 2/5
- cool — 0 — 1 — 1/5

$S.D(temp) = \left[\frac{2}{5} * 3.53\right] + \left[\frac{2}{5} * 9.19\right] + \left[\frac{1}{5} * 0\right]$

$= 5.37$

SDR (temp) = S.D(target) - S.D(temp)

= 8.70 - 5.37

= 3.33

Humidity
- High — 7.071 — 3 — 3/5
- Normal — 7.07 — 2 — 2/5

$S.D(Humidity) = \left[\frac{3}{5} * 7.071\right] + \left[\frac{2}{5} * 7.07\right]$

= 7.07

SDR (Humidity) = S.D(T) - S.D(Humidity)

= 8.70 - 7.07 = 1.63

Windy	False	9.19	3	3/5
	True	12.72	2	2/5

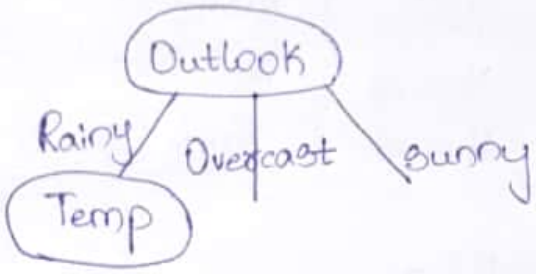
S.D(windy) = $[\frac{3}{5} * 9.19] + [\frac{2}{5} * 12.72]$

= 11.03

SDR = S.D (target) - S.D (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 -

S.D(T) = 4.0311

Mean = 46.25

% C(T) = $\frac{S.D}{mean} \times 100 = 8.71$

Here the % C(T) ≤10, number of samples = 4

So, Here it containg leaf node.

Sunny table :-

Temperature	Humidity	Windy	Hours to play
Mild	High	False	45
Cool	High	False	52
Cool	Normal	True	23
Mild	Normal	False	46
Mild	High	False	30

Target column :-

Mean = 39.2

S.D (target) = 12.15

$\%\,CV = \frac{S.D(T)}{mean} \times 100 = 30.9$

Temperature	Mild	8.96	3	3/5
	Cool	20.30	2	2/5

S.D = 13.57

SDR = 12.15 − 13.576

= −1.42

Humidity	High	11.23	3	3/5
	Normal	16.26	2	2/5

S.D(H) = 13.24

SDR = 12.15 − 13.24

= −1.09

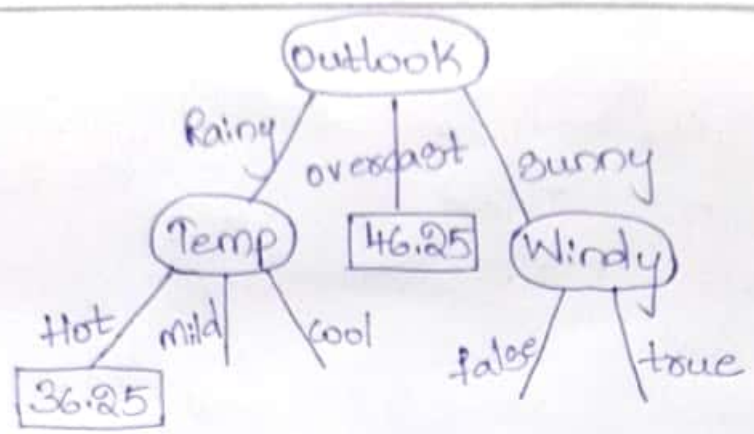
Windy	False	3.78	3	3/5
	True	4.94	2	2/5

S.D(W) = 4.244

SDR = S.D(target) − S.D(W)

= 12.13 − 4.244

= 7.91



Mild table

Humidity	Windy	Hours to play
High	False	45
High	False	35
Normal	False	48
Normal	True	46
High	True	52

Target

mean = 45.2

S.D = 6.370

% CV = $\frac{S.D}{Mean} \times 100$ = 13.928

Humidity → High 8.54 3 3/5

Humidity → Normal 1.41 2 2/5

S.D(Humidity) = 5.688

SDR (Humidity) = S.D(T) − S.D(Humidity)

= 6.30 − 5.688

= 0.62

Cool table :-

Humidity	Windy	Hours to play
Normal	False	52
Normal	True	23
Normal	True	43
Normal	False	38

Mean (T) = 39

S.D(T) = 12.13

% CV = $\frac{S.D}{mean} \times 100 = 31.10$

Humidity Normal 12.13 5/5

S.D(Humidity) = 12.13

SDR(Humidity) = S.D(T) − S.D(Humidity)

= 6.30 − 12.13

= −5.83