

Artificial Intelligence Assignment

Decision tree Regression algorithm.

outlook	temperature	humidity	windy	hours to play
Rainy	Hot	high	False	25
Rainy	Hot	high	true	30
overcast	Hot	High	False	46
Sunny	mild	Normal	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	40
Rainy	mild	Normal	true	48
overcast	mild	high	true	52
overcast	Hot	Normal	False	44
Sunny	mild	high	True	30

criteria = b. / $n=4$

⇒ Decision tree to predict the no. of hours played

Artificial Intelligence Assignment

Decision tree Regression algorithm. . .

outlook	temperature	humidity	windy	hours to play
Rainy	Hot	high	False	25
Rainy	Hot	high	true	30
overcast	Hot	High	False	46
Sunny	mild	Normal	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

criteria = 10% $n=4$

⇒ Decision tree to predict the no. of hours played

Humidity	mean	Standard deviation	count(n)
High	37.57	9.36	7
Normal	42	8.73	7

$$SD(\text{Humidity}) = \frac{7}{14} \times 9.36 + \frac{7}{14} \times 8.73 = 9.05$$

$$SDR = 0.27$$

(=)

temperature	mean	Standard deviation	Count(n)
Hot	27.5	2.5	2
Cool	38	0	1
Mild	41.5	6.5	2

$$SD(\text{temp}) = \frac{2}{5} \times 2.5 + \frac{1}{5} \times 0 + \frac{2}{5} \times 6.5 = 3.6$$

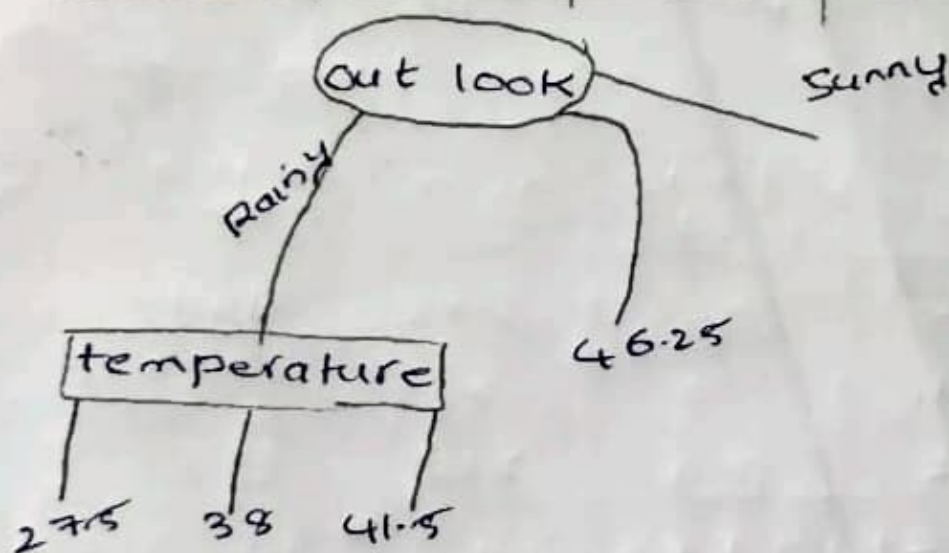
$$SDR = 418$$

Humidity	mean	stand deviation	count (n)
High	30	4.08	3
normal	43.43	5	2

$$S.D(\text{Lumidity}) = \frac{3}{4} \times 4.08 + \frac{2}{5} \times 5 = 4.45$$

$$SDR = 0.35 \times 3.33$$

windy	mean	standard deviation	count(n)
true	29	9	2
False	32.66	5.56	3



=> For outlook sunny!-

temperature	humidity	windy	House played
mild	high	False	45
cool	normal	False	52
cool	normal	true	23
mild	normal	False	46
mild	high	true	30

⇒ Standard deviation of hours played

$$SD(\text{hours}) = 9.32$$

⇒ calc... of Standard deviation of Features Columns

outlook	mean	Stand deviation	count(n)
Sunny	39.2	10.87	5
Rainy	35.2	7.78	5
overcast	45.2	3.32	4

$$\text{outlook} = \frac{5}{14} \times 10.87 + \frac{5}{14} \times 7.78 + \frac{4}{14} \times$$

$$3.49 = 7.66$$

$$SD :- 1.66$$

temper-ature	mean	standard deviat.	count(n)
Hot	36.25	8.8	4
cool	39	10.0	4
overcast	42.66	7.6	6

$$SD(\text{temperature}) = \frac{4}{14} \times 8.95 + \frac{4}{14} \times 10.51 +$$

$$\frac{6}{14} \times 7.65 = 8.84$$

$$SDR = 0.48$$

$$SD(\text{Hours played}) = 10.87$$

$$SD(\text{temperature}) = \frac{3}{5} \times 7.32 + \frac{2}{5} \times 14.15$$

$$= 10.19$$

$$SDR = 0.68$$

humidity	mean	Standard Deviation	count(n)
high	37.5	2.5	2
normal	40.35	12.30	3

$$SD(\text{humidity}) = \frac{2}{5} \times 2.5 + \frac{3}{4} \times 12.5 = 10.50$$

$$SDR = 0.37$$

windy	mean	Standard Dev.	count(n)
true	26.5	3.5	2
false	47.67	3.09	3

$$SD(\text{windy}) = \frac{2}{5} \times 3.5 + \frac{3}{3} \times 3.09$$

$$= 3.25$$

$$SDR = 7.61$$

=> windy has highest SDR it becomes
next node. branch
=> and nodes added with mean
values at output

