Roll No:- 19K41A0517

15-11-2021

ARTIFICIAL INTELLIGENCE

ASSIGNMENT-7

Q) Draw a decision tree diagram to Predict number of hours to play based on weather conditi--ons like outlook, temperature, humidity, windy. Consider dataset given.
(Termination criteria : CV<=10% (or) n=4)

Solution:

Root node identification:

Step:1:-

Standard deviation of Play hours.: SD(Hours) = 9.3210

Step:2:-

Find the standard deviation reduction of the Attributes.

Outlook	Mean	SD	Count.
Sunny.	39.2	10.8701	5
Rainy.	35.2	7.7820	5
Overcast	46.25	3.4910	4

SD = 7.65

SDR = $5/14 * 10.87 + 5/14 * 7.78 + 4/14 * 3.4910$

SDR = 1.662150337

Roll No:- 19K41A0517

Temperature	Mean	SD	count.
Hot	36.25	8.954	4
cool	39	10.511	4
Mild	42.6666	7.652	6.

SD = 8.8413

SDR = $4/14 \times 8.954 + 4/14 \times 10.511 + 6/14 \times 7.652$

SDR = 0.4796.

Humidity	Mean	SD	count.
High.	37.5714	9.3634	7
Normal	42	8.7341	7.

SD = 9.0487

SDR = $14/7 \times 9.3634 + 14/7 \times 8.7341$

SDR = 0.2722

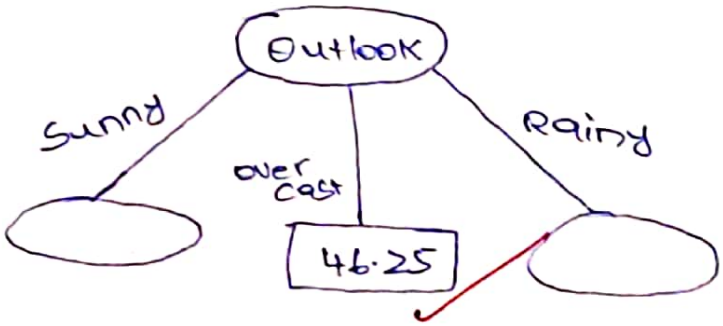
Windy.	Mean	SD	count
TRUE	37.666	10.5934	6
FALSE	41.375	7.8780	8

SD = 9.03893

SDR = 0.2821

The SDR of Outlook Attribute is high so it will be the Root node of decision tree.

Roll NO:- 19K41A0517



Above is the initial structure of decision tree, the overcast has reached decision node since the termination criteria is matched i.e. CV ≤ 10% and n = 4.

leaf node identification:

SD (Hours) = 10.8701426.

outlook : Sunny :

Temperature.	Mean	SD	Count.
Cool	37.5	14.5	2
Mild.	40.33	7.3181	3

SD = 10.19089

SDR = 0.6792.

Humidity.	Mean	SD	Count
High	37.5	7.5	2
Normal	40.33	12.4988.	3

SD = 10.4993

SDR = 0.37080

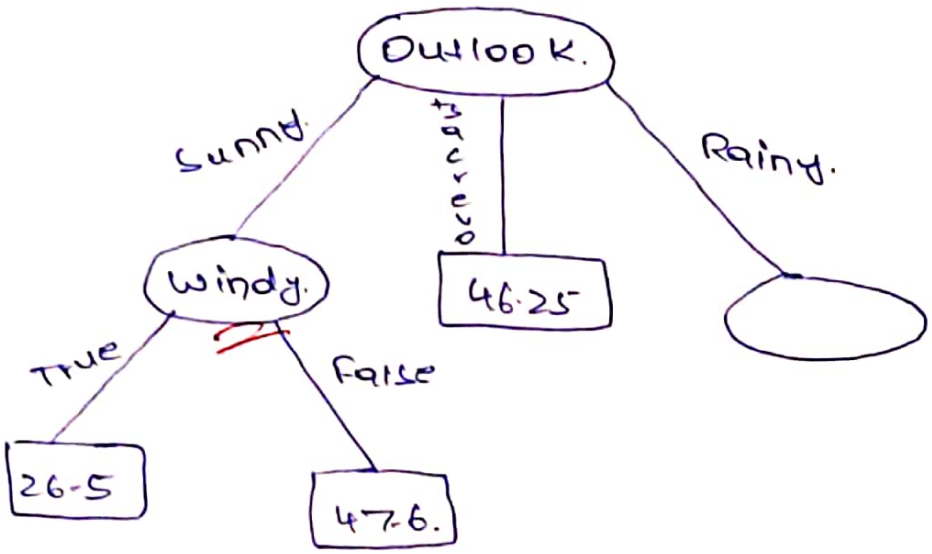
Roll NO:- 19K41A0517

Windy.	Mean.	SD	Count
True.	26.5.	3.5	2
False.	47.66	3.09	3

SD = 3.25472

SDR = 7.6154

The SDR of Windy is higher than the other attributes so, it will be the decision leaf node. 1.



The windy attribute has two leaf nodes which reach the termination criteria of n=4 and CV ≤ 10%

Outlook : Rainy!

Temperature	Mean	SD	count.
Hot	27.5	2.5	2
Cool	38	0	1
mild.	41.5	6.5	2

Roll NO:- 19K41A0517

SD = 3.6.

SDR = 4.182030

Humidity.	Mean	SD	Count
High	30	4.0824	3
Normal	43	5	2.

SD = 4.4494

SDR = 3.33254

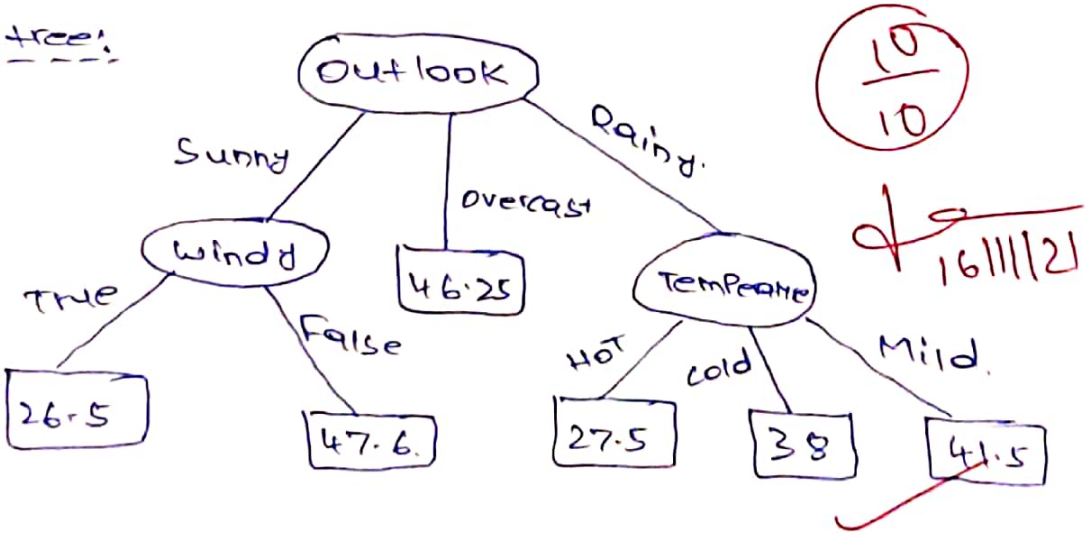
Windy.	Mean	SD	Count
True	39	9	2
false.	32.666	5.5577	3

SD = 6.934 6664

SDR = 0.847

The SDR of temperature is high so, it will be the other decision node.

Decision tree:



Above, is the Decision tree which has all the leaf nodes meeting termination criteria (n=4, CV≤10%