ARTIFICIAL INTELLIGENCE

ASSIGNMENT-7

1) Draw a decision tree diagram to Predict number of hours to Play based on weather conditions like outlooks, temPerature, humidity, windy consider dataset given.

(Termination criteria: CVC=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
overact	46.25	3.4910	4.

SD:= 7.65 SDR = 5/14 * 10.87 + 5/14 * 7.78 + 4/14 * 3.4910

SDR = 1.662150337.

Temparature	Mean	SD	count
Hot	3625	8.954	4
cool	39	10.511	4
mild	42.6666	7.652	6

SD : 8.8413 SDR : 4/14 * 8.954 + 4/14 * 10.51 + 6/14 * 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 * 9.3634 + 14/7 * 8.7341

SDR = 0.2722

Windy

Windy	mean	SD	count
True	37.666	10.5934	6
FALSE	41.375	7.870	8

SDR = 0.2824

The SDR of outlook Attribute is high so it will be the root node of decision tree.

Above is the intial structure of decision tree the overcast has reached decision node since the termination criterion is matched i.e

$CV \leq 10.1$ and $n = 4$.

leaf node identification:

SD (Hours) = 10.87014426.

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	12.4987	3

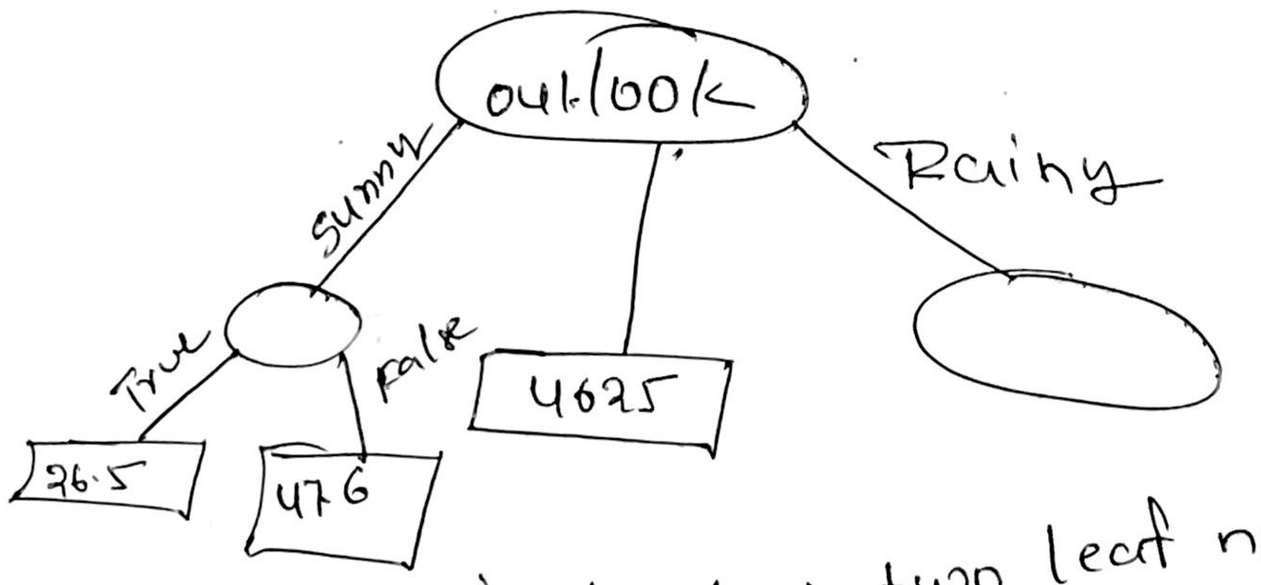
SD = 10.4993 SDR = 0.37080

Windy	mean	SD	count
True	26.5	3.5	2
False	47.66	3.09	3

SD = 3.25472

SDR = 7.6454

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



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

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

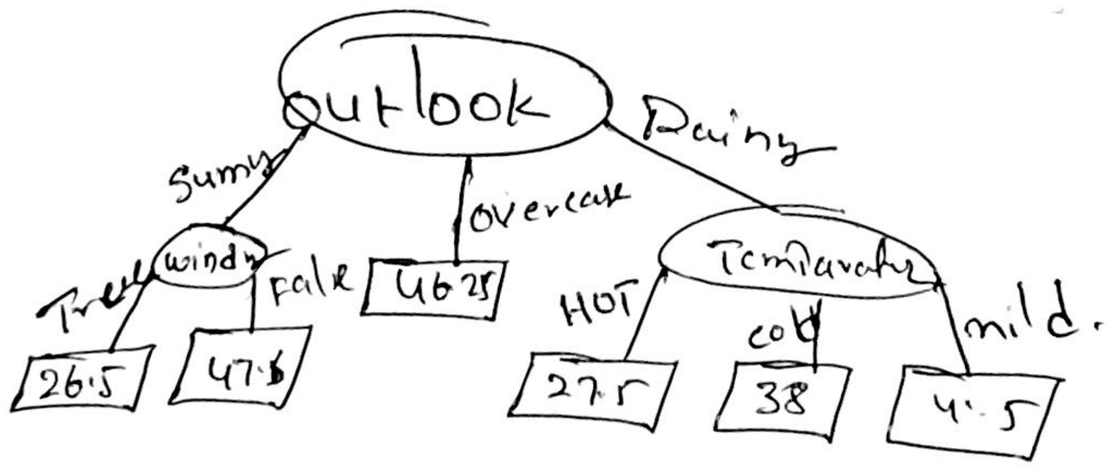
Windy	mean	SD	count
True	39	9	2
False	32.666	5.5577	3

SD = 6.9346664

SDR = 0.847

The SDR of temparature 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,c=10).