SUB- ARTIFICIAL INTELLIGENCE Roll. No: 19KUIA0516

ASSIGNMENT - 7

1) Draw a decision tree diagram to predict numb of hours to stay based on weather condition like outlook, temperature, humidity windy consider the dataset given
(Termination Criteria : CV <= 10% or n = 4)

Ans: Step-1:

Standard deviation of play hours SD(Hours) = 9.3210

Step-2:

To find the standard deviation reduction of attributes

Outlook	Mean	SD	Sum
Sunny	39.2	10.8701	5
Rainy	35.2	7.7820	5
Overcase	46.25	3.4910	4

SD = 7.65

SDD = 5% * 10.87 + 5% * 7.78 + 4% * 3.49

SDD = 1.662150337

ROLL NO: 19K41A0516

Temperature	Mean	SD	Count
Hot	36.25	8.954	4
cool	3.9	10.511	4
mild	42.6666	7.652	6

SD = 8.8413 SDD = 4/14 * 8.95 + 4/14 * 10.511 + 6/14 * 7.652

SDD = 0.4796

Humidity	Mean	SD	Count
High	37.5714	9.3634	7
Normal	42	8.7344	7

SD = 9.0487 SDD = 7/14 * 9.8634 + 7/14 * 6.7344

SDD = 0.2722

Windy	Mean	SD	Count
True	32.666	10.5934	6
False	41.375	7.5730	8

SD = 9.03893 SDR = 0.2821

The SDR value of Outlook attributes is high so so it will be the root node of decision tree.

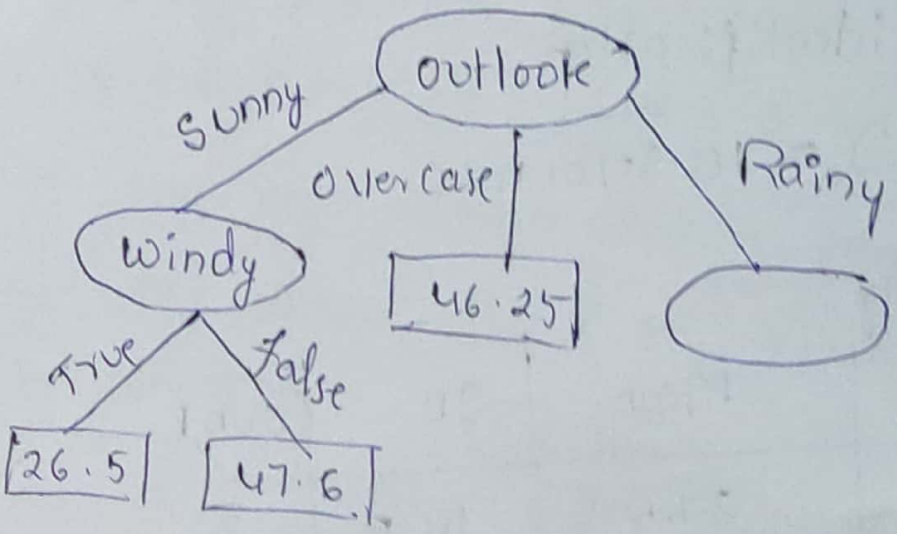
Windy

Roll No: 19K41A0516

Windy	Mean	SD	Count
true	26.5	3.8	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 decision leaf node



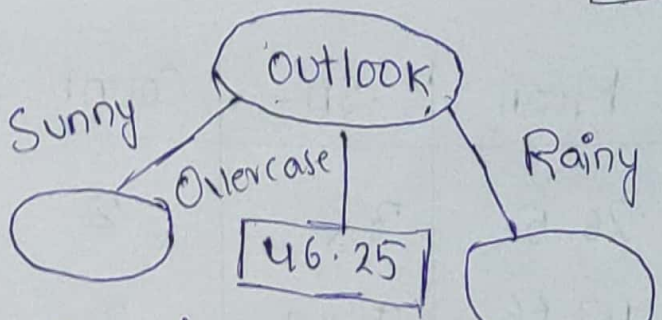
The windy attribute has two leaf nodes which reach the termination criteria of n=4 and CV ≤ 10%

Outlook Rainy;

Temperature	Mean	SD	Count
Hot	27.8	2.5	2
cool	38	0	1
mild	41.5	6.5	2

SD = 3.6 | SDR = 4.1820 |

Roll No: 19K41



Above is the initial attribute of decision tree the Overcase has reached the decision node. Since the termination criteria is matched i.e CV ≤ 10% 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	12.49	3

SD = 10.4993 | SDR 0.37050

Roll no : 19KU1A0516

Humidity	Mean	SD	Count
High	3.0	4.08	3
Normal	4.3	5	2

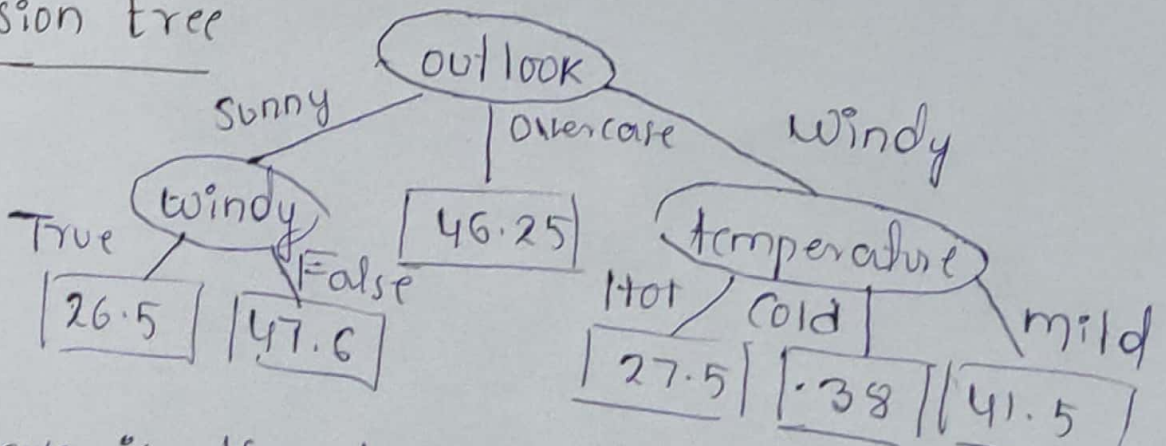
SD = 4.494 | SDR = 3.33234

Windy	Mean	SD	Count
True	39	9	2
False	32.666	5.5577	3

SD = 6.934664 | SDR = 0.847

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

Decision tree



→ Above is the decision tree which has attributes for which termination criteria n=4 and cv ≤ 10%