Decision Tree assignment:

Outlook (y(9)/N(5))

Sunny (y(2)/N(3)) | Overcast (y(4)/N(0)) Pure | Rain (y(3)/N(2))

~~Gain~~

Using the function for gain & entropy provided in the Jupyter notebook, and doing calculations,

gain = (entropy(9,5), [5/14, 4/14, 5/14], [entropy(2/3), entropy(3/2)])

= 0.24

Temp. (y(9)/N(5))

Hot (y(2)/N(2)) | mild (y(4)/N(2)) | cool (y(3)/N(1))

gain = (entropy(9,5), [4/14, 6/14, 4/14], [entropy(2,2), entropy(4,2), entropy(3,1)]

= 0.03

Humidity (9/5)

High (y(3)/N(4)) | Normal (y(6)/N(1))

gain = (entropy(9,5), [7/14, 7/14], [e(3,4), e(6,1)]

= 0.15

Wind (y(9)/N(5))

Weak (y(6)/N(2)) | Strong (y(3)/N(3))

gain = (entropy(9,5), [8/14, 6/14], [entropy(6,2), e(3,3)])

= 0.04

Now, the "outlook" is the root node, as these ~~case~~ is max gain in it.

For the root node, Sunny,

Temp. (y(2)/N(3))

Hot (0/2) | Mild (1/1) | Cool (1/0)

Gain = (entropy(2,3), [2/5, 2/5, 1/5], [0, e(1,1), 0]

= -0.02

Humidity (2/3)

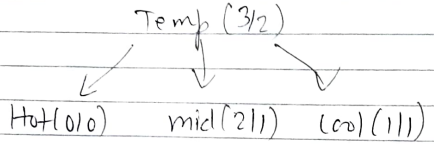
High (0/3) | Normal (2,0)

gain = (entropy (2,3), [3/5, 2/5], [0,0])

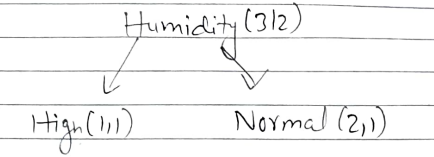
= 0.97

Since, the gain is max in humidity, the second root node becomes "Sunny".

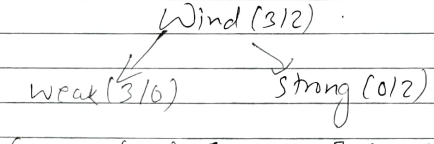
Now, for Rain,



Gain = (entropy(3,2), [3/5, 2/5], [entropy(2,1), entropy(1,1)]
= 0.01



Gain = (entropy(3,2), [2/5, 3/5], [entropy (1,1), entropy (2,1)]
= 0.01



Gain = (entropy (3/2), [3/5, 2/5], [0,0]
= 0.97

Therefore, the wind is the root node for this case as it has the maximum value.

Hence, the final decision tree is presented below: