

Q2. Handling missing values in decision trees: Assume that the Outlook of D3 is missing. Could you still learn a decision tree from this example? The set of features of interest are: Outlook, Temperature, Humidity and Wind. Clarifying again - Day is not to be used as a feature in this sub part of the question. Show the learnt decision tree despite the missing Outlook values from D3.

Ans:

As we can find out from figure 1 that outlook has the highest information gain so it can be the root node for the tree. Now for missing values in the day, we will assign some probability according to the values present to that particular column at the same level in the tree. By calculating all the information gain, I can implement a decision tree. When looking for the missing value, it can figure out by the decision tree.

