Decision Tree classifier

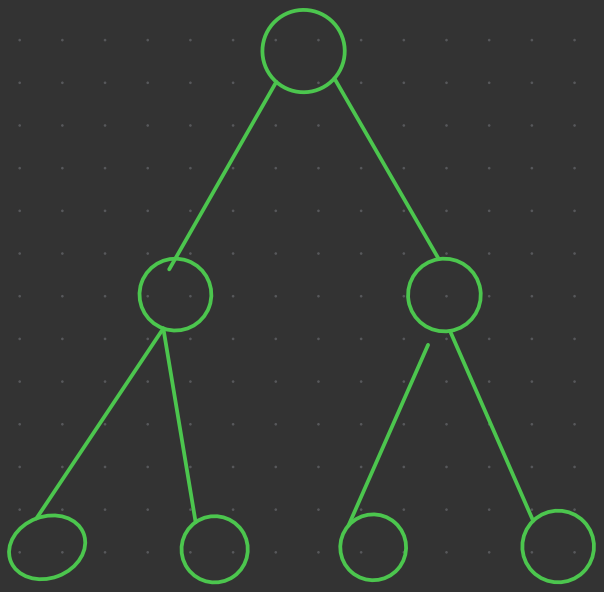
Decision Tree classifier → ID3
Decision Tree classifier → CART

a) Entropy and Gini Index → Purity split
b) Information Gain → Features to select for DT construction

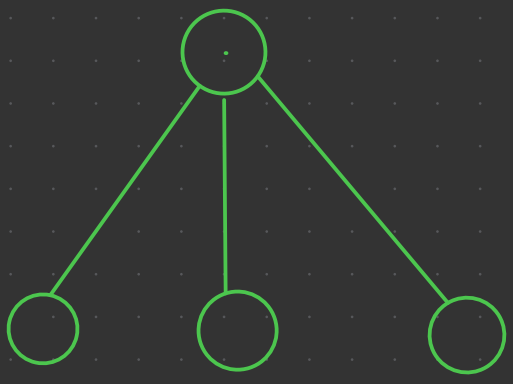
Sklearn uses CART

The difference between ID3 and CART is that, when we use you create decision tree CART create binary splits against each and every node, where as in ID3, we can have several binary splits under one node.

CART



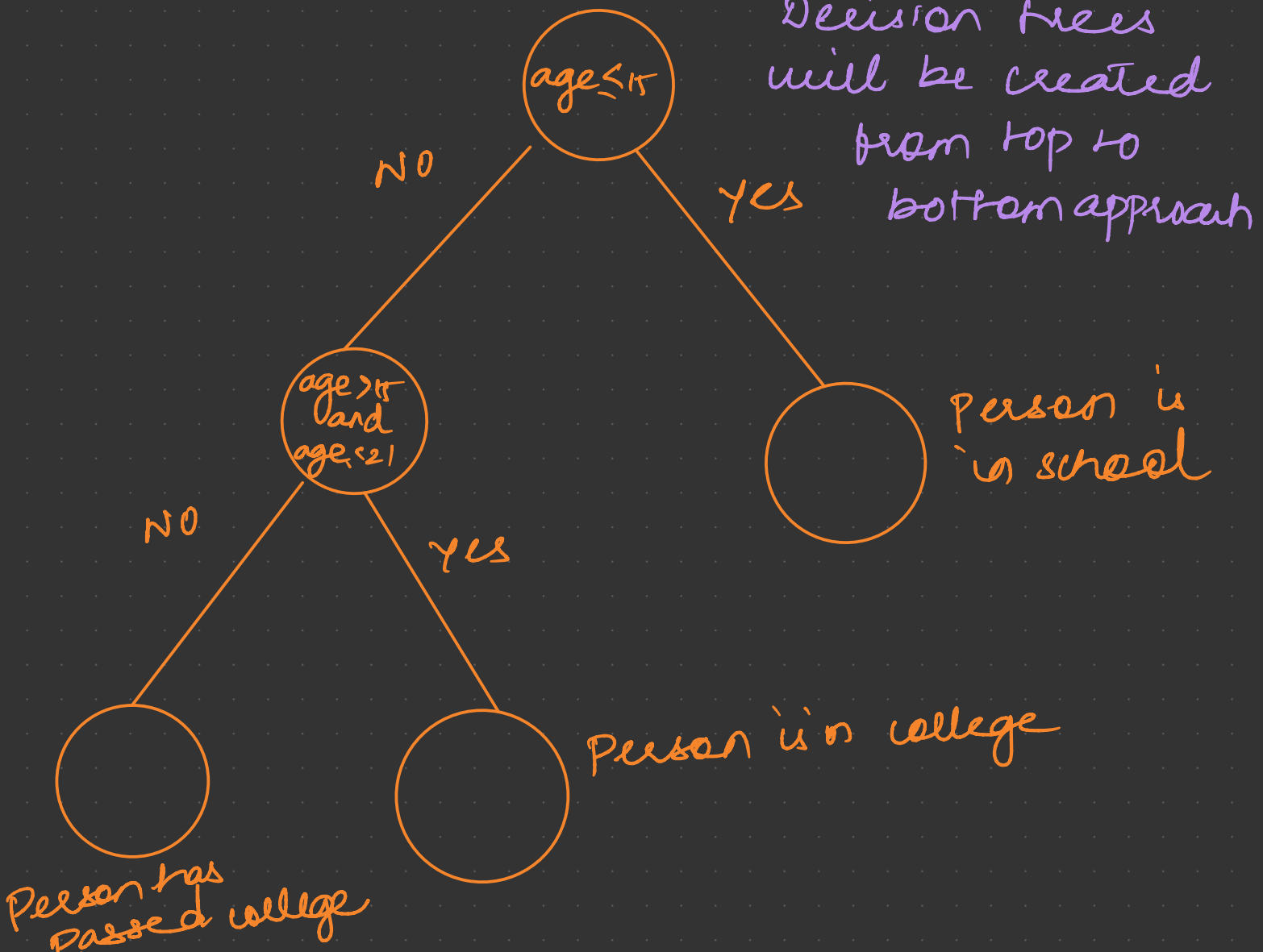
ID3



Above representation clearly says how the binary split will happen internally.

example

```
age = 14
if (age <= 15):
    print ("The person is in school")
elif (age > 15 and age <= 21):
    print ("The person is in college")
else:
    print ("The person has passed college")
```



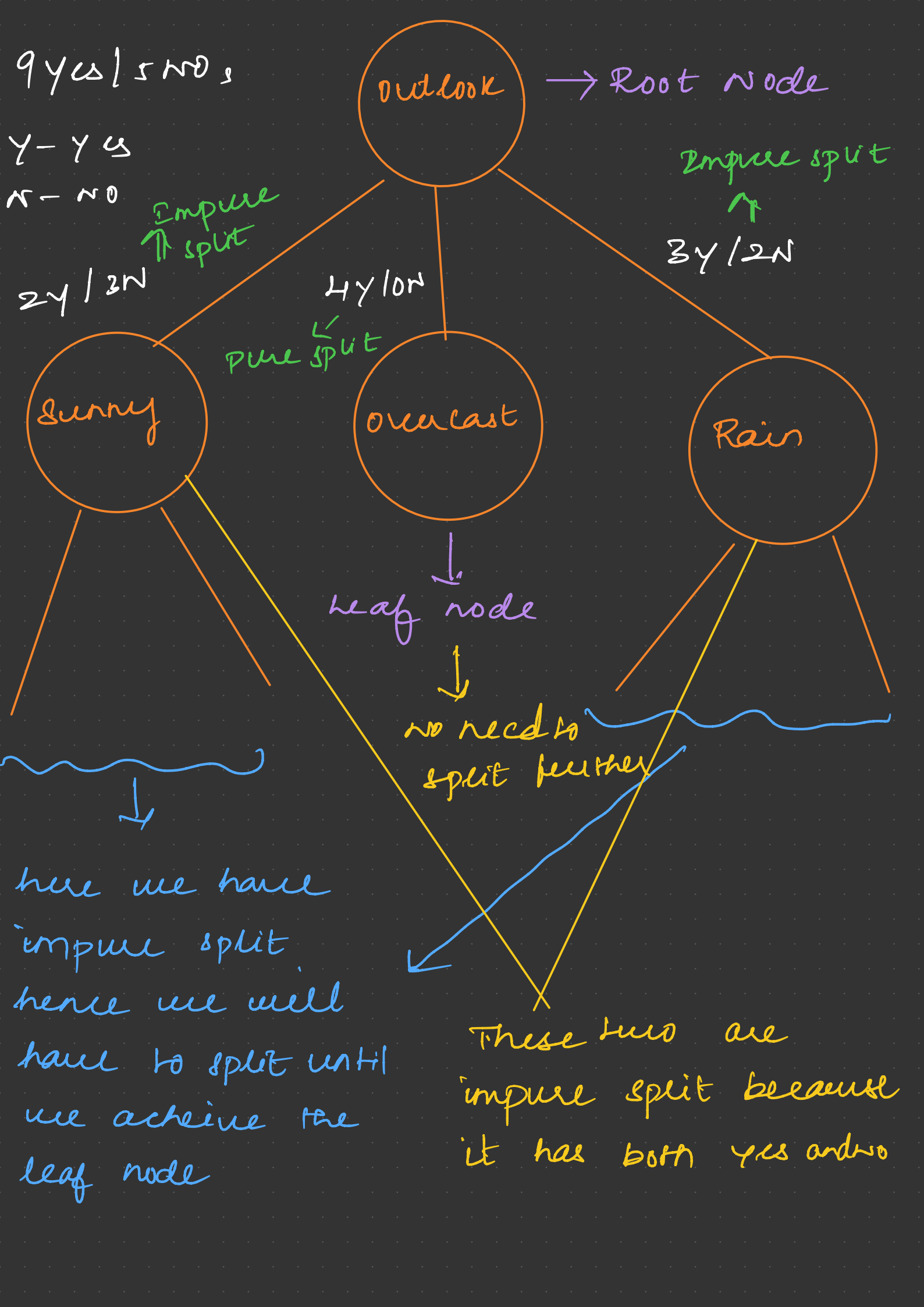
Decision trees will be created from top to bottom approach

Dataset

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

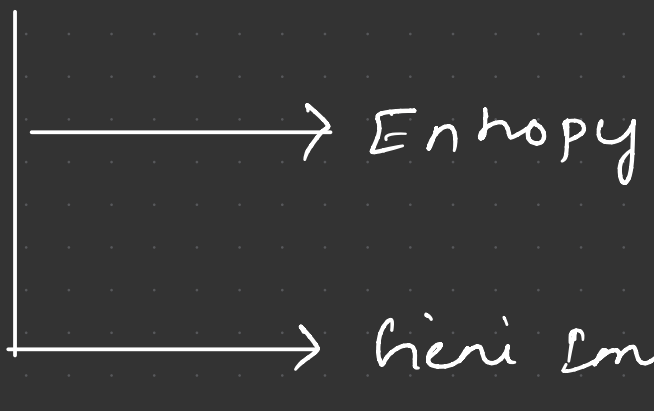
here we have the independent features such as outlook, temprature, humidity, wind, and we really need to predict if they can play tennis or not based on the independent features.

this is obviously a classification problem — "Binary classification" as the dependent output variable is "Yes or NO".


9 yes | 5 no's
Y - Yes
N - No
Outlook
→ Root Node
Impure split
3Y / 2N
Impure split
2Y / 3N
4Y / 0N
Pure split
Sunny
Overcast
Rain
leaf node
no need to split further
here we have impure split hence we will have to split until we achieve the leaf node
These two are impure split because it has both yes and no

how we mathematically find out if the split is pure split or impure split - For that purpose we will be using Entropy and information gain.

① Purity → Pure or Impure



② what feature needs to be selected for split

For this we use Information gain