Decision Trees

Outline

- Decision Tree Representations
  - ID3 learning algorithms (Quinlan 1986)
- Entropy, Information Gain

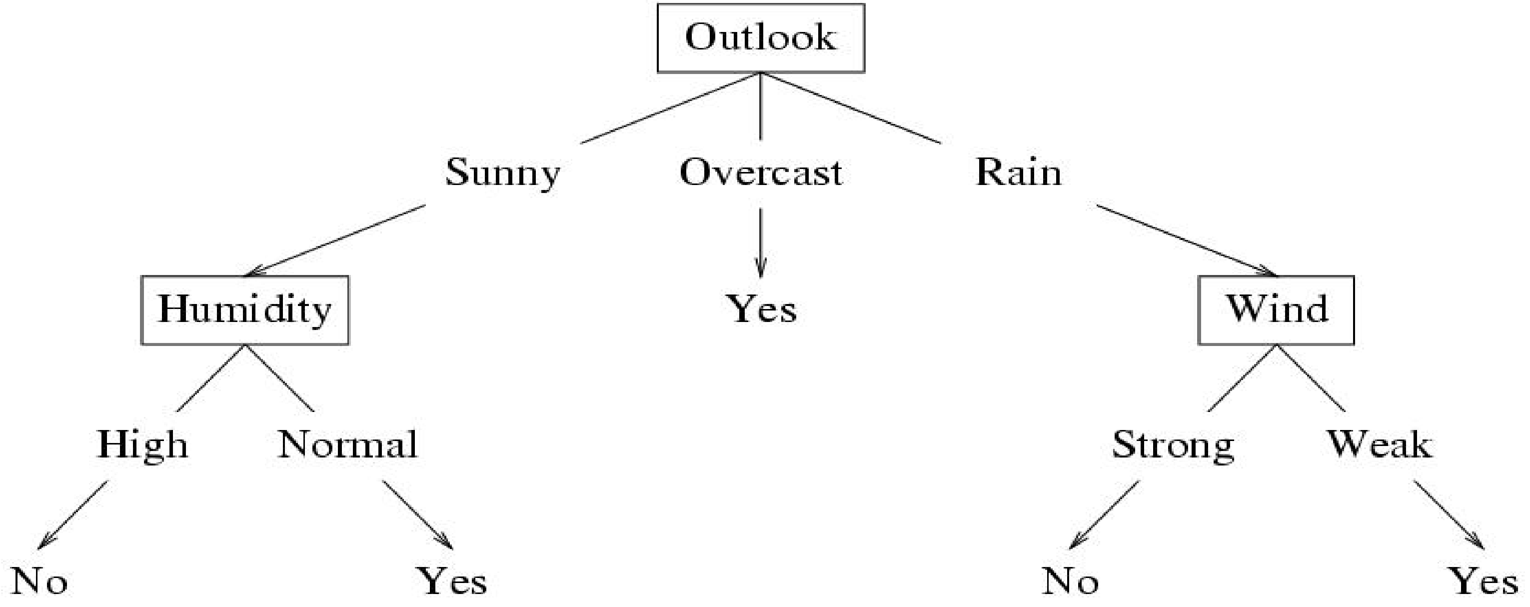
Training Data Example: Goal is to Predict When This Player Will Play Tennis?

PlayTennis: training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

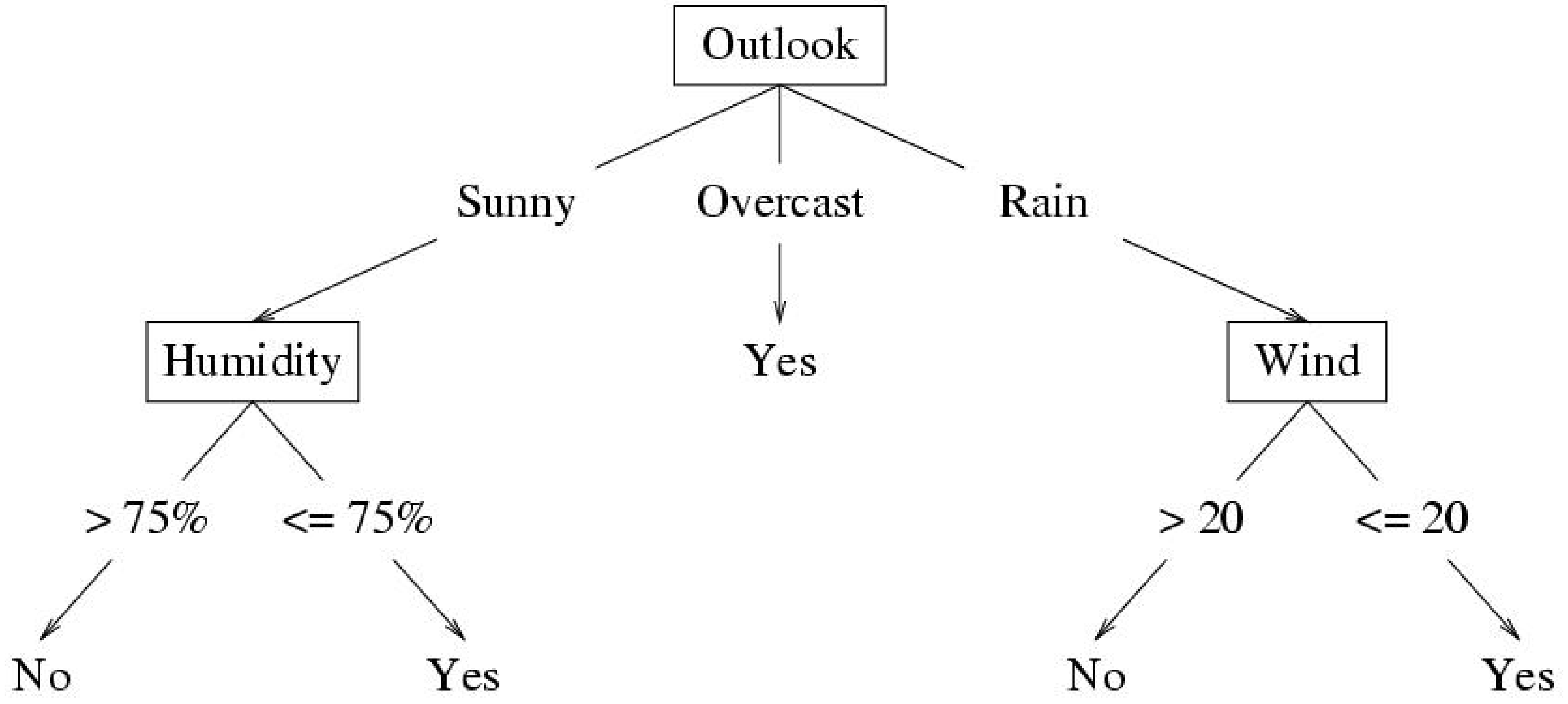
Decision Tree Hypothesis Space

- **Internal nodes** test the value of particular features x_j and branch according to the results of the test.
- **Leaf nodes** specify the class $h(\mathbf{x})$.

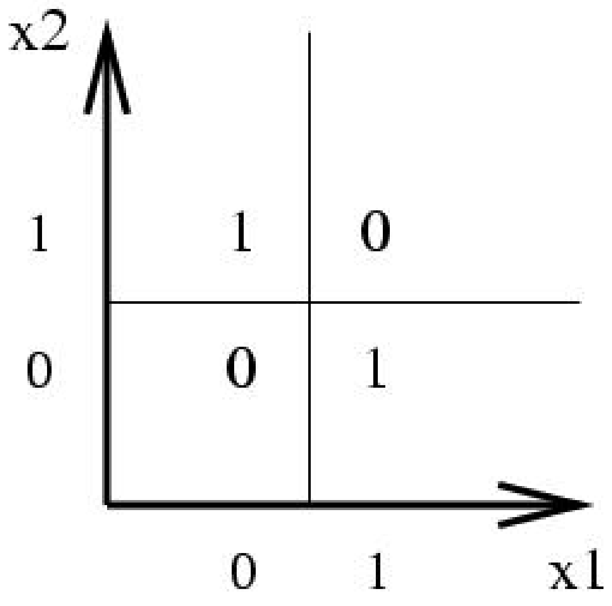


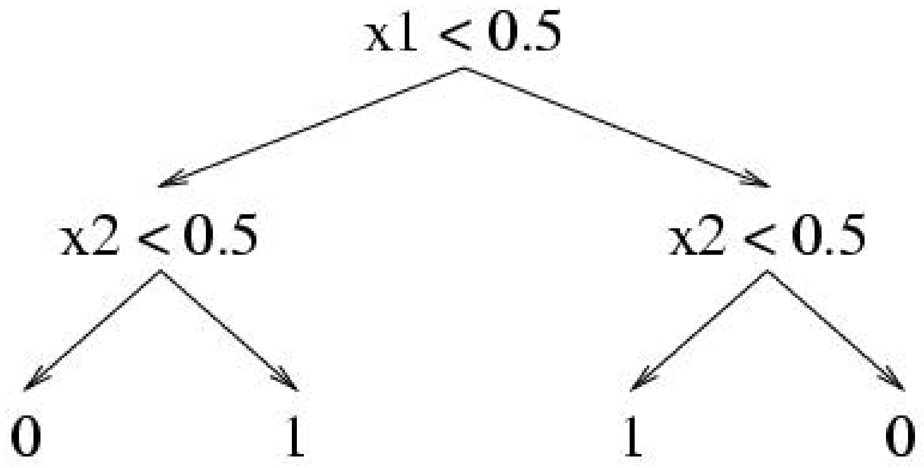
Suppose the features are **Outlook** (x_1), **Temperature** (x_2), **Humidity** (x_3), and **Wind** (x_4). Then the feature vector $\mathbf{x} = (Sunny, Hot, High, Strong)$ will be classified as **No**. The **Temperature** feature is irrelevant.

If the features are continuous, internal nodes may test the value of a feature against a threshold.



Decision Trees Can Represent Any Boolean Function



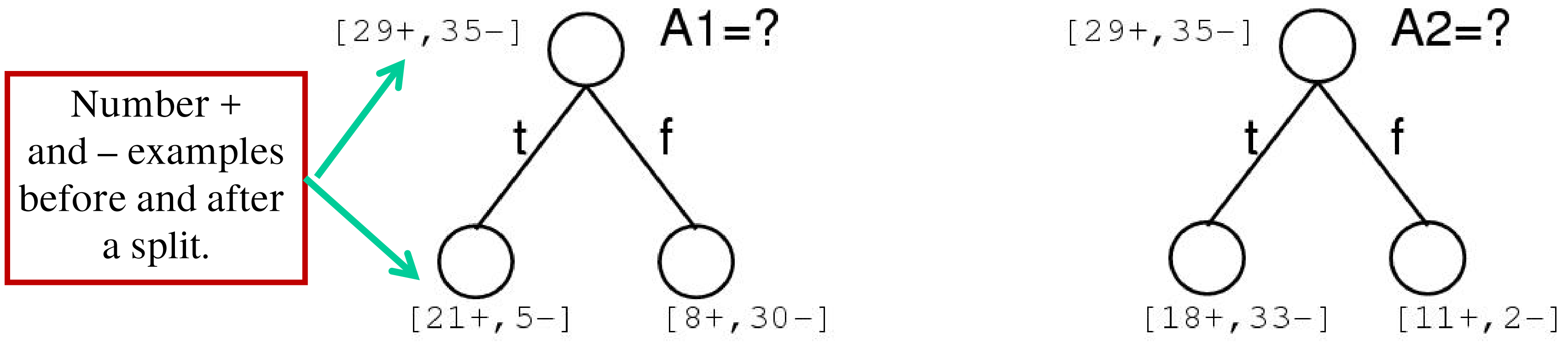


The tree will in the worst case require exponentially many nodes, however.

Choosing the Best Attribute

A1 and A2 are "attributes" (i.e. features or inputs).

Which attribute is best?



- Many different frameworks for choosing BEST have been proposed!
- We will look at Entropy Gain.

Entropy

- Entropy is a concept used in decision tree algorithms to measure the impurity or randomness of a set of examples within a specific class.

- In the context of decision trees, entropy is commonly used to determine the best attribute to split the data on at each node.

- $p_{\oplus}$ is the proportion of positive examples in S
- $p_{\ominus}$ is the proportion of negative examples in S
- Entropy measures the impurity of S

$$Entropy(S) \equiv -p_{\oplus} \log_2 p_{\oplus} - p_{\ominus} \log_2 p_{\ominus}$$

Entropy

- Entropy measures the randomness/uncertainty in the data
- Let's consider a set S of examples with C many classes. Entropy of this set:

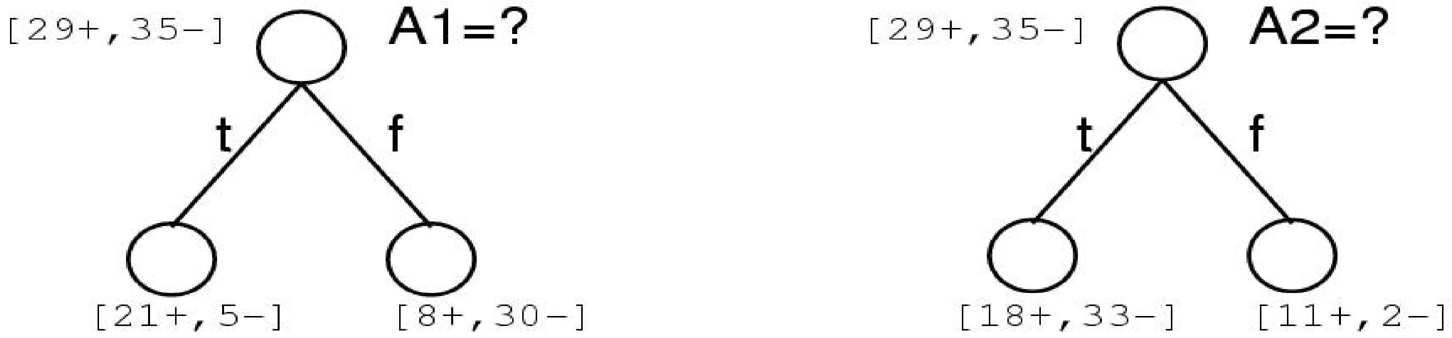
$$H(S) = -\sum_{c \in C} p_c \log_2 p_c$$

- p_c is the probability that an element of S belongs to class c
  - .. basically, the fraction of elements of S belonging to class c
- Intuition: Entropy is a measure of the "degree of surprise"
  - Some dominant classes $\Longrightarrow$ small entropy (less uncertainty)
  - Equiprobable classes $\Longrightarrow$ high entropy (more uncertainty)
- Entropy denotes the average number of bits needed to encode S

Information Gain

$Gain(S, A)$ = expected reduction in entropy due to sorting on A

$$Gain(S, A) \equiv Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$



Information Gain

- Let's assume each element of S consists of a set of features
- Information Gain (IG) on a feature F

$$IG(S, F) = H(S) - \sum_{f \in F} \frac{|S_f|}{|S|} H(S_f)$$

- S_f number of elements of S with feature F having *value* f
- $IG(S, F)$ measures the increase in our certainty about S once we know the value of F
- $IG(S, F)$ denotes the number of bits saved while encoding S once we know the value of the feature F

Computing Information Gain

- Let's begin with the root node of the DT and compute IG of each feature
- Consider feature "wind" $\in$ {weak,strong} and its IG w.r.t. the root node

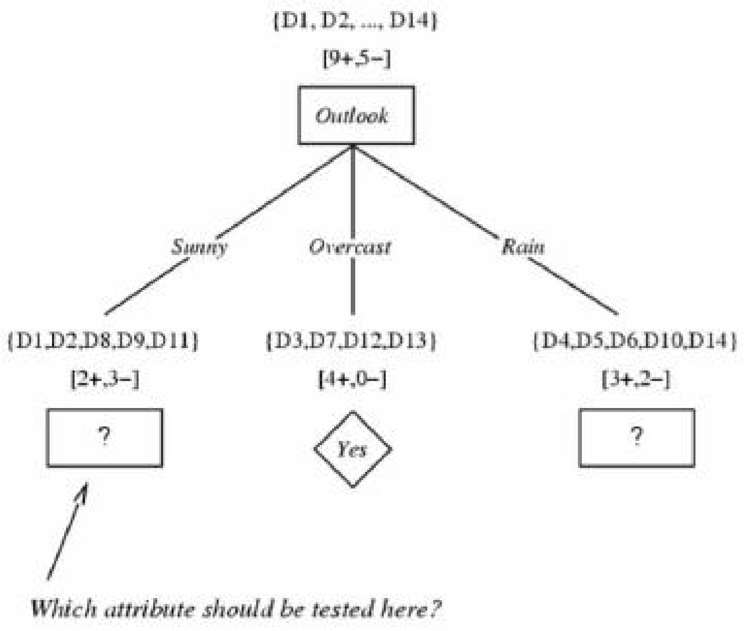
day	outlook	temperature	humidity	wind	play
1	sunny	hot	high	weak	no
2	sunny	hot	high	strong	no
3	overcast	hot	high	weak	yes
4	rain	mild	high	weak	yes
5	rain	cool	normal	weak	yes
6	rain	cool	normal	strong	no
7	overcast	cool	normal	strong	yes
8	sunny	mild	high	weak	no
9	sunny	cool	normal	weak	yes
10	rain	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

- Root node: $S = [9+, 5-]$ (all training data: 9 play, 5 no-play)
- Entropy: $H(S) = -(9/14)\log_2(9/14) - (5/14)\log_2(5/14) = 0.94$
- $S_{weak} = [6+, 2-] \Longrightarrow H(S_{weak}) = 0.811$
- $S_{strong} = [3+, 3-] \Longrightarrow H(S_{strong}) = 1$

$$\begin{aligned} IG(S, \text{wind}) &= H(S) - \frac{|S_{weak}|}{|S|} H(S_{weak}) - \frac{|S_{strong}|}{|S|} H(S_{strong}) \\ &= 0.94 - 8/14 * 0.811 - 6/14 * 1 \\ &= 0.048 \end{aligned}$$

Choosing the most informative feature

- At the root node, the information gains are:
  - $IG(S, \text{wind}) = 0.048$ (we already saw)
  - $IG(S, \text{outlook}) = 0.246$
  - $IG(S, \text{humidity}) = 0.151$
  - $IG(S, \text{temperature}) = 0.029$
- "outlook" has the maximum $IG \Longrightarrow$ chosen as the root node
- Growing the tree:
  - Iteratively select the feature with the highest information gain for each child of the previous node



Example

PlayTennis: training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Attribute: Outlook

Values (Outlook) = Sunny, Overcast, Rain

$S = [+9, -5]$ $\quad Entropy(S) = -\frac{9}{14}log_2\frac{9}{14} - \frac{5}{14}log_2\frac{5}{14} = 0.94$

$S_{sunny} = [+2, -3]$ $\quad Entropy(S_{sunny}) = -\frac{2}{5}log_2\frac{2}{5} - \frac{3}{5}log_2\frac{3}{5} = 0.971$

$S_{Overcast} = [+4, 0]$ $\quad Entropy(S_{Overcast}) = -\frac{4}{4}log_2\frac{4}{4} - \frac{0}{4}log_2\frac{0}{4} = 0$

$S_{Rain} = [+3, -2]$ $\quad Entropy(S_{Rain}) = -\frac{3}{5}log_2\frac{3}{5} - \frac{2}{5}log_2\frac{2}{5} = 0.971$

$$Gain(S, Outlook) = Entropy(S) - \sum_{v \in (sunny, overcas,\ rain)} \frac{|S_v|}{|s|} Entropy(s_v)$$

$Gain(S, Outlook) =$

$= Entropy(S) - \frac{5}{14}Entropy(S_{sunny}) - \frac{4}{14}Entropy(S_{Overcast}) - \frac{5}{14}Entropy(S_{Rain})$

$Gain(S, Outlook) = 0.94 - \frac{5}{14}*0.971 - \frac{4}{14} * 0 - \frac{5}{14} * 0.971 = 0.2464$

PlayTennis: training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Attribute: Temperature

Values (Temperature) = Hot, Mild, Cool

$S = [+9, -5]$ $\quad Entropy(S) = -\frac{9}{14}log_2\frac{9}{14} - \frac{5}{14}log_2\frac{5}{14} = 0.94$

$S_{Hot} = [+2, -2]$ $\quad Entropy(S_{Hot}) = -\frac{2}{4}log_2\frac{2}{4} - \frac{2}{4}log_2\frac{2}{4} = 1.0$

$S_{Mild} = [+4, -2]$ $\quad Entropy(S_{Mild}) = -\frac{4}{6}log_2\frac{4}{6} - \frac{2}{6}log_2\frac{2}{6} = 0.9183$

$S_{Cool} = [+3, -1]$ $\quad Entropy(S_{Cool}) = -\frac{3}{4}log_2\frac{3}{4} - \frac{1}{4}log_2\frac{1}{4} = 0.8113$

$$Gain(S, Temperature) = Entropy(S) - \sum_{v\in(Hot,Mild,Cool)} \frac{|S_v|}{|s|} Entropy(s_v)$$

$Gain(S, Temperature) =$

$= Entropy(S) - \frac{4}{14}Entropy(S_{Hot}) - \frac{6}{14}Entropy(S_{Mild}) - \frac{4}{14}Entropy(S_{Cool})$

$Gain(S, Temperature) = 0.94 - \frac{4}{14}*1.0 - \frac{6}{14} * 0.9183 - \frac{4}{14} * 0.8113 = 0.0289$

PlayTennis: training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

Attribute: Humidity

Values (Humidity) = High, Normal

Similarly ……

$$Gain(S, Humidity) = 0.94 - \frac{7}{14} * 0.9852 - \frac{7}{14} * 0.5916 = 0.1516$$

Attribute: Wind

Values (Wind) = Strong, Weak

Similarly ……

$$Gain(S, Wind) = 0.94 - \frac{6}{14} * 1.0 - \frac{8}{14} * 0.8113 = 0.0478$$

PlayTennis: training examples

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

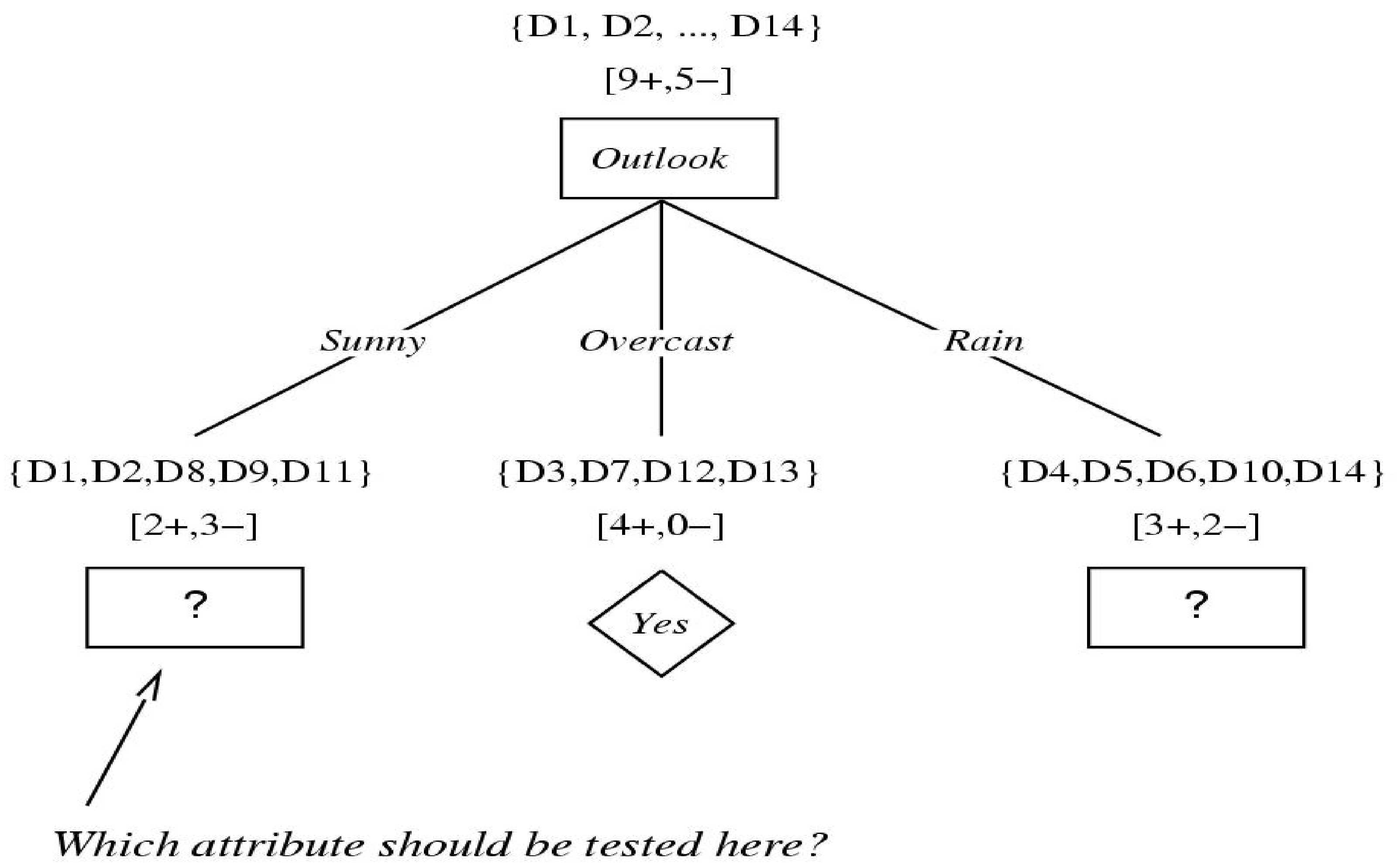
Gain(S, Outlook) = 0.2464

Gain(S, Temperature) 0.0289

Gain(S, Humidity) = 0.1516

Gain(S, Wind) = 0.0478

Outlook **has the maximum Gain.**
Thus, it should be the root node in the tree.



S_{sunny} = {D1,D2,D8,D9,D11}

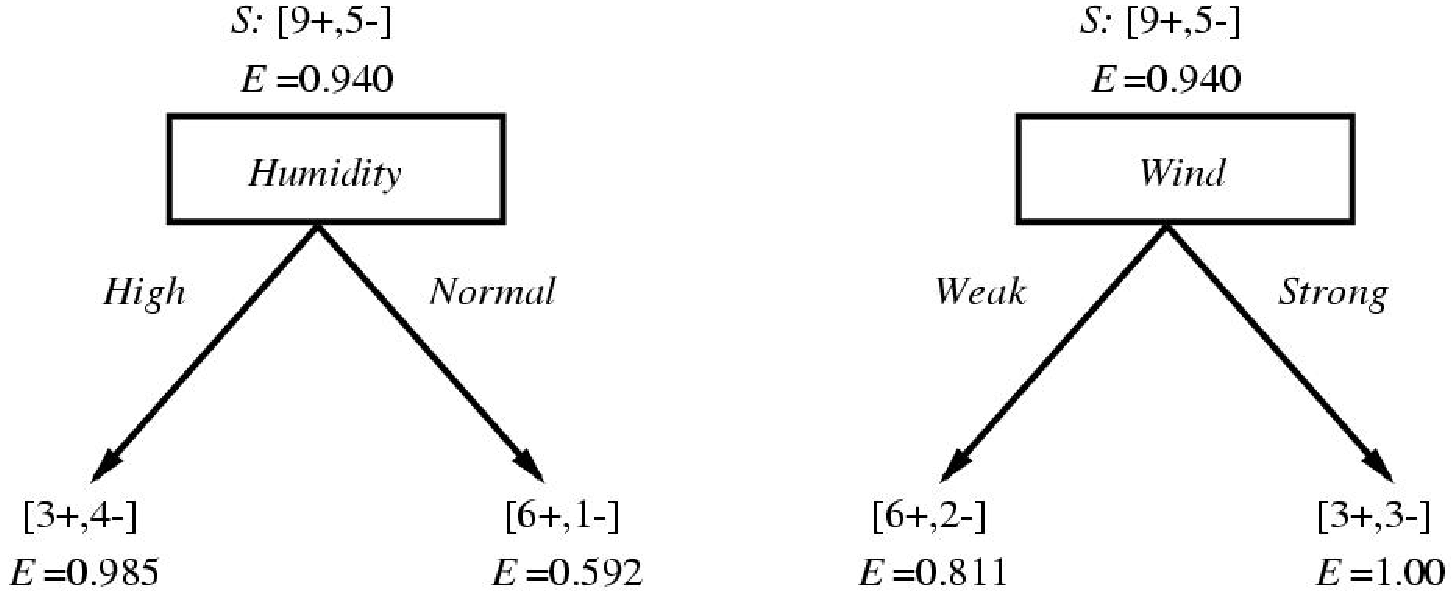
$Gain\ (S_{sunny}, Humidity) = .970 - (3/5)\ 0.0 - (2/5)\ 0.0 = .970$

$Gain\ (S_{sunny}, Temperature) = .970 - (2/5)\ 0.0 - (2/5)\ 1.0 - (1/5)\ 0.0 = .570$

$Gain\ (S_{sunny}, Wind) = .970 - (2/5)\ 1.0 - (3/5)\ .918 = .019$

Selecting the Next Attribute

Which attribute is the best classifier?



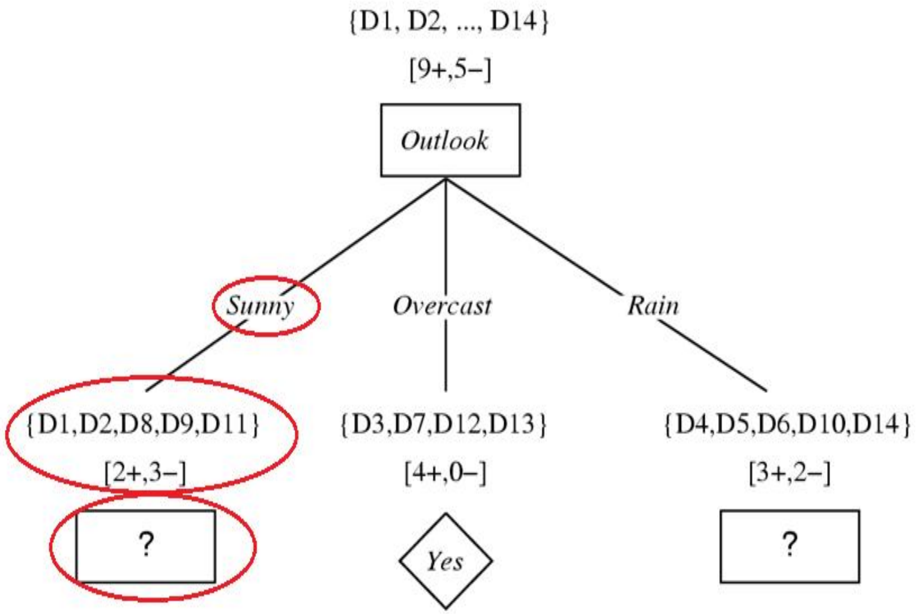
Gain (S, Humidity)

= .940 - (7/14).985 - (7/14).592
= .151

Gain (S, Wind)

= .940 - (8/14).811 - (6/14)1.0
= .048

Day	Temperature	Humidity	Wind	Play Tennis
D1	Hot	High	Weak	No
D2	Hot	High	Strong	No
D8	Mild	High	Weak	No
D9	Cool	Normal	Weak	Yes
D11	Mild	Normal	Strong	Yes



Attribute: Temperature

Values (Temperature) = Hot, Mild, Cool

$S_{sunny} = [+2, -3]$ $Entropy(S) = -\frac{2}{5}log_2\frac{2}{5} - \frac{3}{5}log_2\frac{3}{5} = 0.97$

$S_{Hot} = [0, -2]$ $Entropy(S_{Hot}) = -\frac{0}{2}log_2\frac{0}{2} - \frac{2}{2}log_2\frac{2}{2} = 0$

$S_{Mild} = [+1, -1]$ $Entropy(S_{Mild}) = -\frac{1}{2}log_2\frac{1}{2} - \frac{1}{2}log_2\frac{1}{2} = 1$

$S_{Cool} = [+1, 0]$ $Entropy(S_{Cool}) = 0$

$$Gain(S_{sunny}, Temperature) = Entropy(S) - \sum_{v\in(Hot,Mild,Cool)} \frac{|S_v|}{|S|} Entropy(s_v)$$

$$Gain(S_{sunny}, Temperature) =$$
$$= Entropy(S) - \frac{2}{5}Entropy(S_{Hot}) - \frac{2}{5}Entropy(S_{Mild}) - \frac{1}{5}Entropy(S_{Cool})$$

$$Gain(S_{sunny}, Temperature) = 0.97 - \frac{2}{5}*0.0 - \frac{2}{5} * 1 - \frac{1}{5} * 0.0 = 0.570$$

Day	Temperature	Humidity	Wind	Play Tennis
D1	Hot	High	Weak	No
D2	Hot	High	Strong	No
D8	Mild	High	Weak	No
D9	Cool	Normal	Weak	Yes
D11	Mild	Normal	Strong	Yes

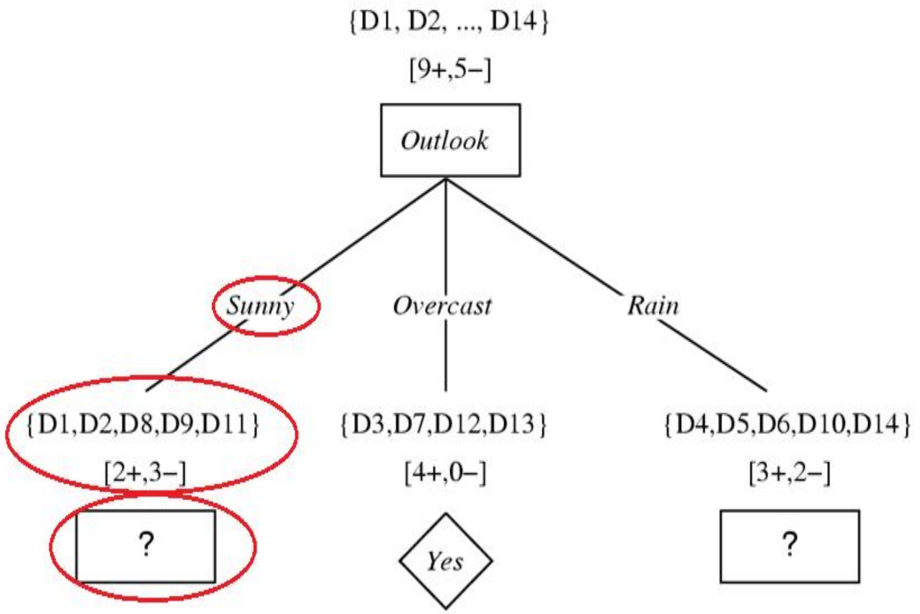
Attribute: Humidity

Values (Humidity) = High, Normal

S_{sunny} = [+2, -3] $Entropy(S) = -\frac{2}{5}log_2\frac{2}{5} - \frac{3}{5}log_2\frac{3}{5} = 0.97$

S_{High} = [0, -3] $Entropy(S_{High}) = 0$

S_{Normal} = [+2, -0] $Entropy(S_{Mild}) = 0$

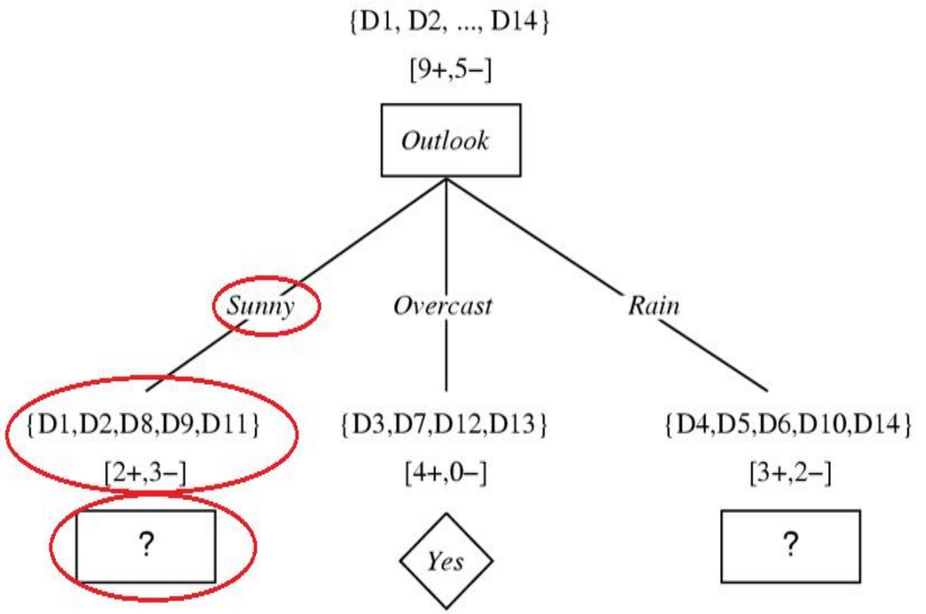


$$Gain(S_{sunny}, Humidity) = Entropy(S) - \sum_{v\in(High,\ Normal)} \frac{|S_v|}{|s|} Entropy(s_v)$$

$$Gain(S_{sunny}, Humidity) =$$
$$= Entropy(S) - \frac{3}{5}Entropy(S_{High}) - \frac{2}{5}Entropy(S_{Normal})$$

$$Gain(S_{sunny}, Humidity) = 0.97 - \frac{3}{5}*0.0 - \frac{2}{5}*0.0 = 0.97$$

Day	Temperature	Humidity	Wind	Play Tennis
D1	Hot	High	Weak	No
D2	Hot	High	Strong	No
D8	Mild	High	Weak	No
D9	Cool	Normal	Weak	Yes
D11	Mild	Normal	Strong	Yes



Attribute: Wind

Values (Wind) = Strong, Weak

S_{sunny} = [+2, -3] $Entropy(S) = -\frac{2}{5}log_2\frac{2}{5} - \frac{3}{5}log_2\frac{3}{5} = 0.97$

S_{Strong} = [+1, -1] $Entropy(S_{strong}) = 1.0$

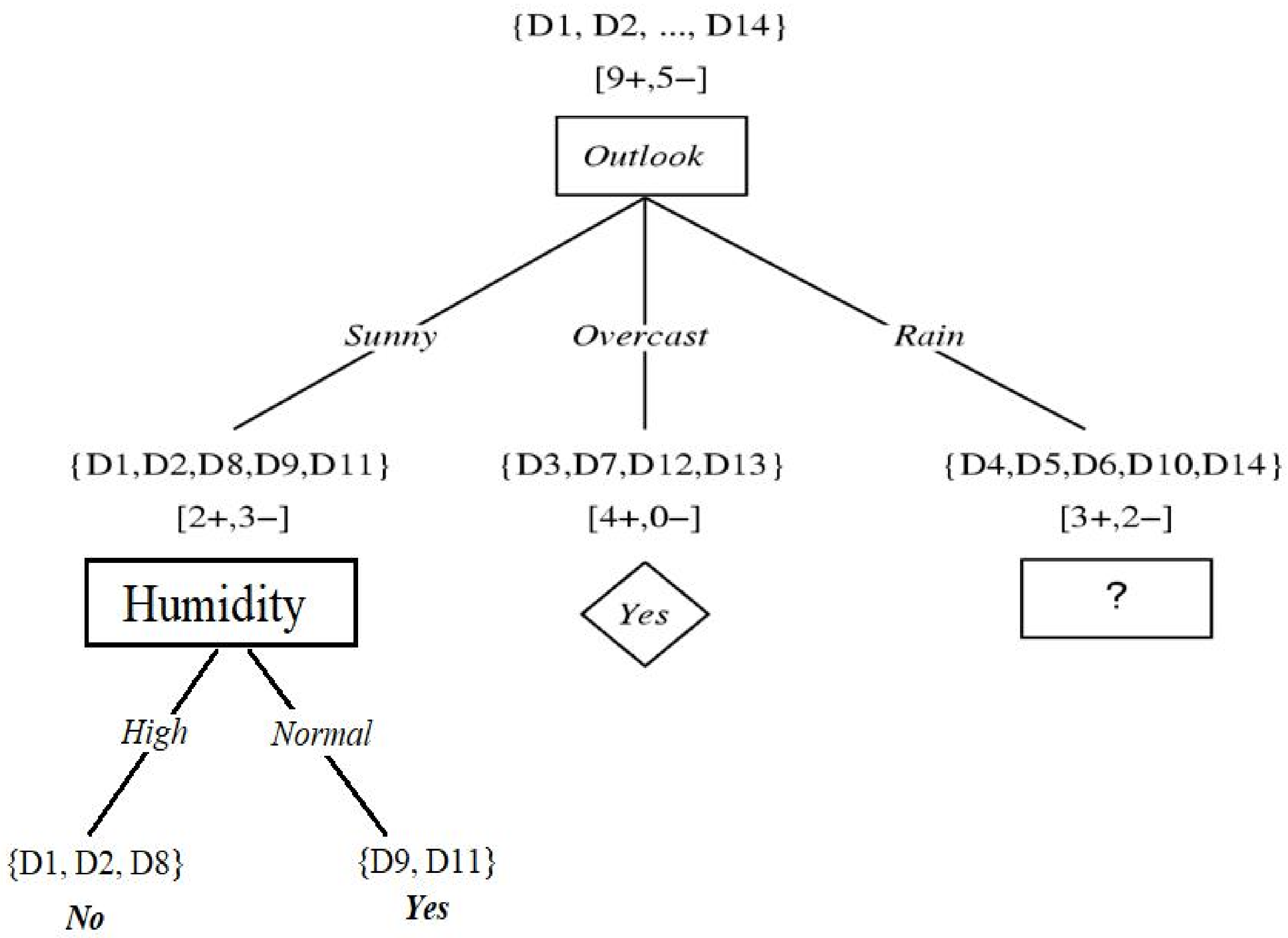
S_{weak} = [+1, -2] $Entropy(S_{weak}) = -\frac{1}{3}log_2\frac{1}{3} - \frac{2}{3}log_2\frac{2}{3} = 0.9183$

$$Gain(S_{sunny}, Wind) = Entropy(S) - \sum_{v \in (Strong,\, Weak)} \frac{|S_v|}{|s|} Entropy(s_v)$$

$$Gain(S_{sunny}, Wind) = Entropy(S) - \frac{2}{5}Entropy(S_{Strong}) - \frac{3}{5}Entropy(S_{Weak})$$

$$Gain(S_{sunny}, Wind) = 0.97 - \frac{2}{5}*1.0 - \frac{3}{5} * 0.918 = 0.0192$$

Day	Temperature	Humidity	Wind	Play Tennis
D1	Hot	High	Weak	No
D2	Hot	High	Strong	No
D8	Mild	High	Weak	No
D9	Cool	Normal	Weak	Yes
D11	Mild	Normal	Strong	Yes

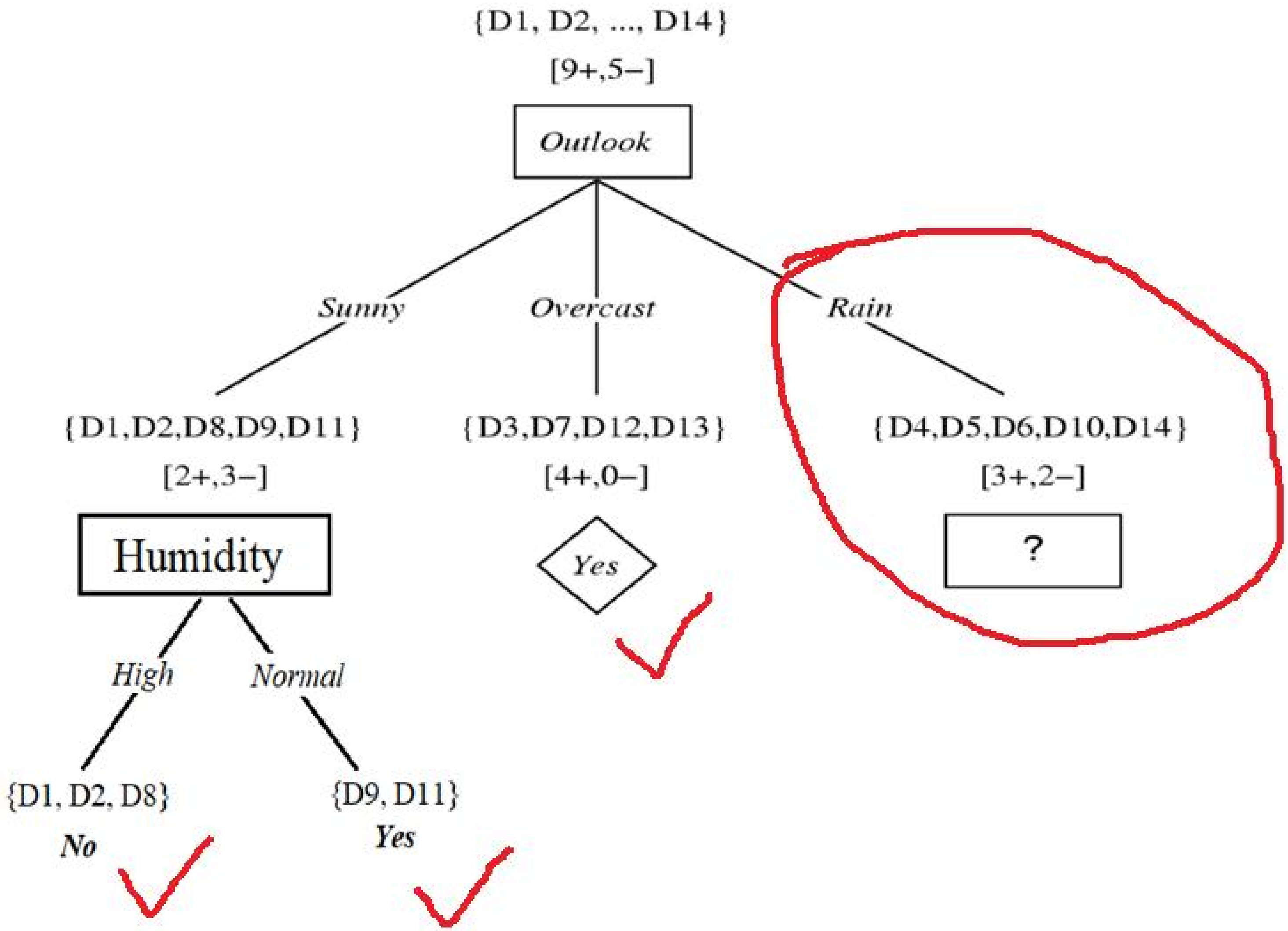


$Gain(S_{sunny}, Temperature) = 0.570$

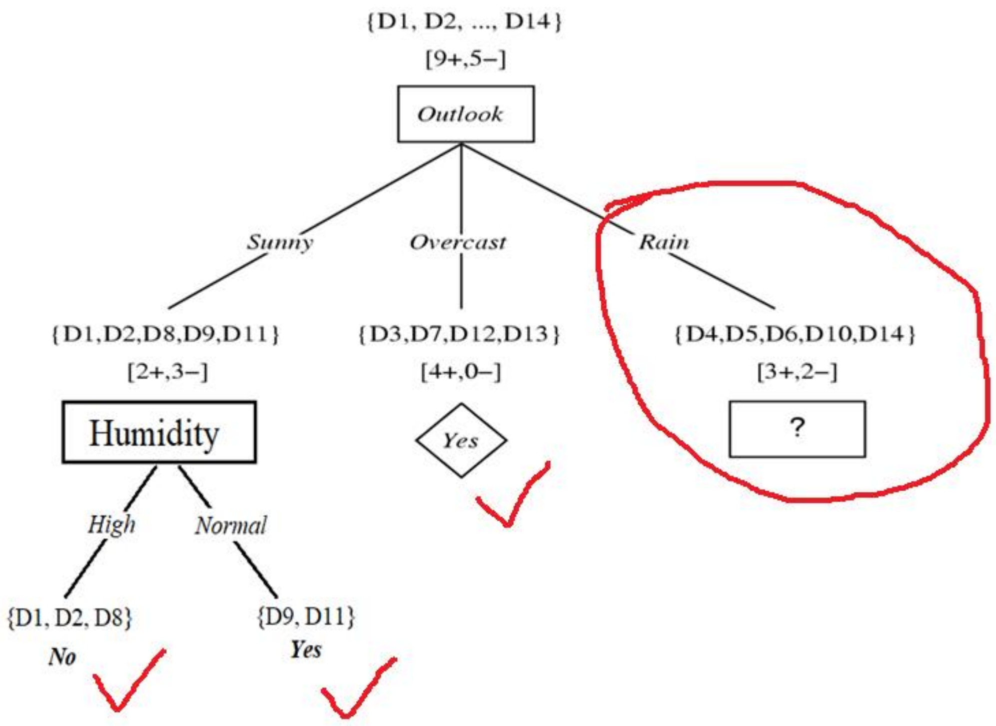
$\underline{Gain(S_{sunny}, Humidity) = 0.97}$

$Gain(S_{sunny}, Wind) = 0.0192$

Day	Temperature	Humidity	Wind	Play Tennis
D4	Mild	High	Weak	Yes
D5	Cool	Normal	Weak	Yes
D6	Cool	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No



Day	Temperature	Humidity	Wind	Play Tennis
D4	Mild	High	Weak	Yes
D5	Cool	Normal	Weak	Yes
D6	Cool	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No



Attribute: Temperature

Values (Temperature) = Hot, Mild, Cool

$S_{Rain} = [+3, -2]$ $\quad$ $Entropy(S) = -\frac{3}{5}log_2\frac{3}{5} - \frac{2}{5}log_2\frac{2}{5} = 0.97$

$S_{Hot} = [+0, -0]$ $\quad$ $Entropy(S_{Hot}) = 0$

$S_{Mild} = [+2, -1]$ $\quad$ $Entropy(S_{Mild}) = -\frac{2}{3}log_2\frac{2}{3} - \frac{1}{3}log_2\frac{1}{3} = 0.9183$

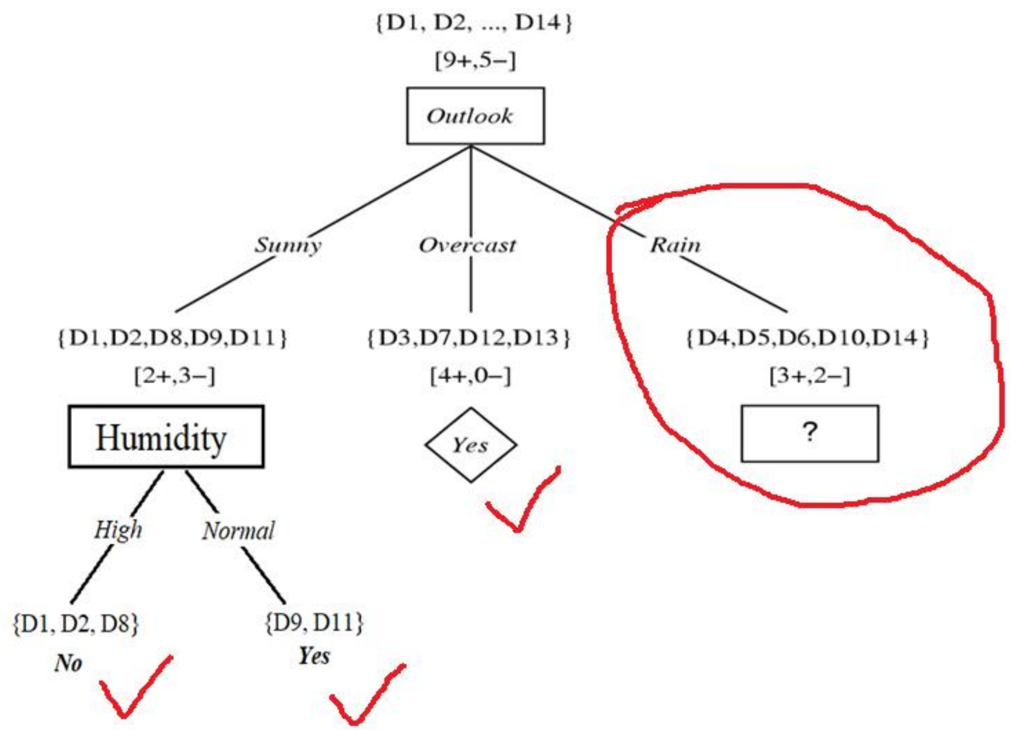
$S_{Cool} = [+1, -1]$ $\quad$ $Entropy(S_{Cool}) = 1.0$

$$Gain(S_{Rain}, Temperature) = Entropy(S) - \sum_{v \in (Hot,\ Mild,\ Cool)} \frac{|S_v|}{|S|} Entropy(s_v)$$

$$Gain(S_{Rain}, Temperature) =$$
$$= Entropy(S) - \frac{0}{5}Entropy(S_{Hot}) - \frac{3}{5}Entropy(S_{Mild}) - \frac{3}{5}Entropy(S_{Cool})$$

$$Gain(S_{Rain}, Temperature) = 0.97 - \frac{0}{5}*0.0 - \frac{3}{5}*0.918 - \frac{3}{5}*1.0 = 0.0192$$

Day	Temperature	Humidity	Wind	Play Tennis
D4	Mild	High	Weak	Yes
D5	Cool	Normal	Weak	Yes
D6	Cool	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No



Attribute: Humidity

Values (Humidity) = High, Normal

S_{Rain} = [+3, -2] $Entropy(S) = -\frac{3}{5}log_2\frac{3}{5} - \frac{2}{5}log_2\frac{2}{5} = 0.97$

S_{High} = [+1, -1] $Entropy(S_{High}) = 1.0$

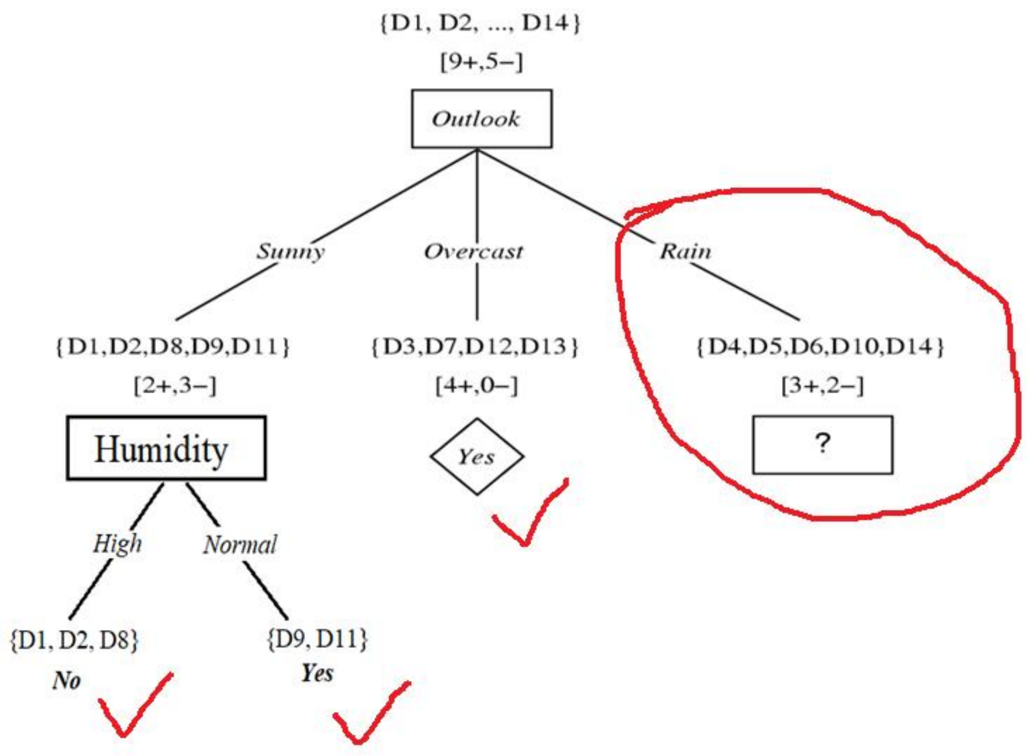
S_{Normal} = [+2, -1] $Entropy(S_{Mild}) = -\frac{2}{3}log_2\frac{2}{3} - \frac{1}{3}log_2\frac{1}{3} = 0.9183$

$$Gain(S_{Rain}, Humidity) = Entropy(S) - \sum_{v\in(High,\ Normal)} \frac{|S_v|}{|s|} Entropy(s_v)$$

$$Gain(S_{Rain}, Humidity) = Entropy(S) - \frac{2}{5} Entropy(S_{High}) - \frac{2}{5} Entropy(S_{Normal})$$

$$Gain(S_{Rain}, Humidity) = 0.97 - \frac{2}{5}*1.0 - \frac{2}{5} * 0.918 = 0.0192$$

Day	Temperature	Humidity	Wind	Play Tennis
D4	Mild	High	Weak	Yes
D5	Cool	Normal	Weak	Yes
D6	Cool	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No



Attribute: Wind

Values (Wind) = Weak, Strong

$S_{Rain} = [+3, -2]$ $\quad$ $Entropy(S) = -\frac{3}{5} log_2 \frac{3}{5} - \frac{2}{5} log_2 \frac{2}{5} = 0.97$

$S_{Weak} = [+3, -0]$ $\quad$ $Entropy(S_{Weak}) = 0.0$

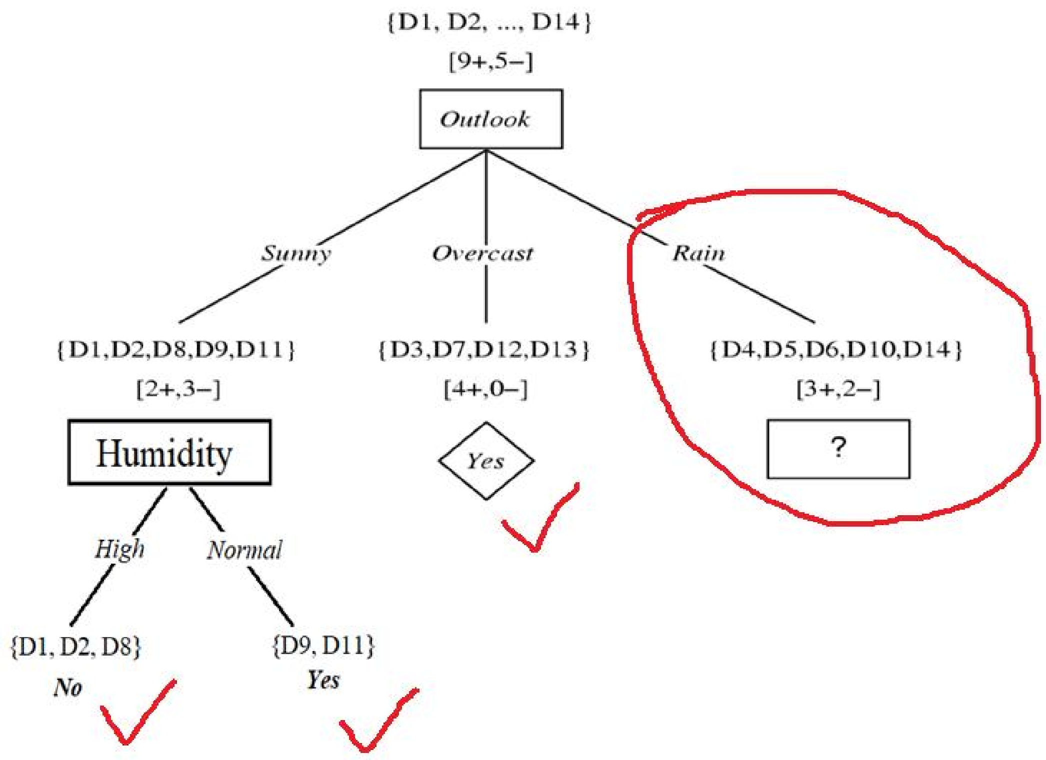
$S_{Strong} = [+0, -2]$ $\quad$ $Entropy(S_{Strong}) = 0.0$

$$Gain(S_{Rain}, Wind) = Entropy(S) - \sum_{v \in (Weak,\ Strong)} \frac{|S_v|}{|S|} Entropy(S_v)$$

$$Gain(S_{Rain}, Wind) = Entropy(S) - \frac{3}{5} Entropy(S_{Weak}) - \frac{2}{5} Entropy(S_{Strong})$$

$$Gain(S_{Rain}, Wind) = 0.97 - \frac{3}{5}*0.0 - \frac{2}{5} * 0.0 = 0.97$$

Day	Temperature	Humidity	Wind	Play Tennis
D4	Mild	High	Weak	Yes
D5	Cool	Normal	Weak	Yes
D6	Cool	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No

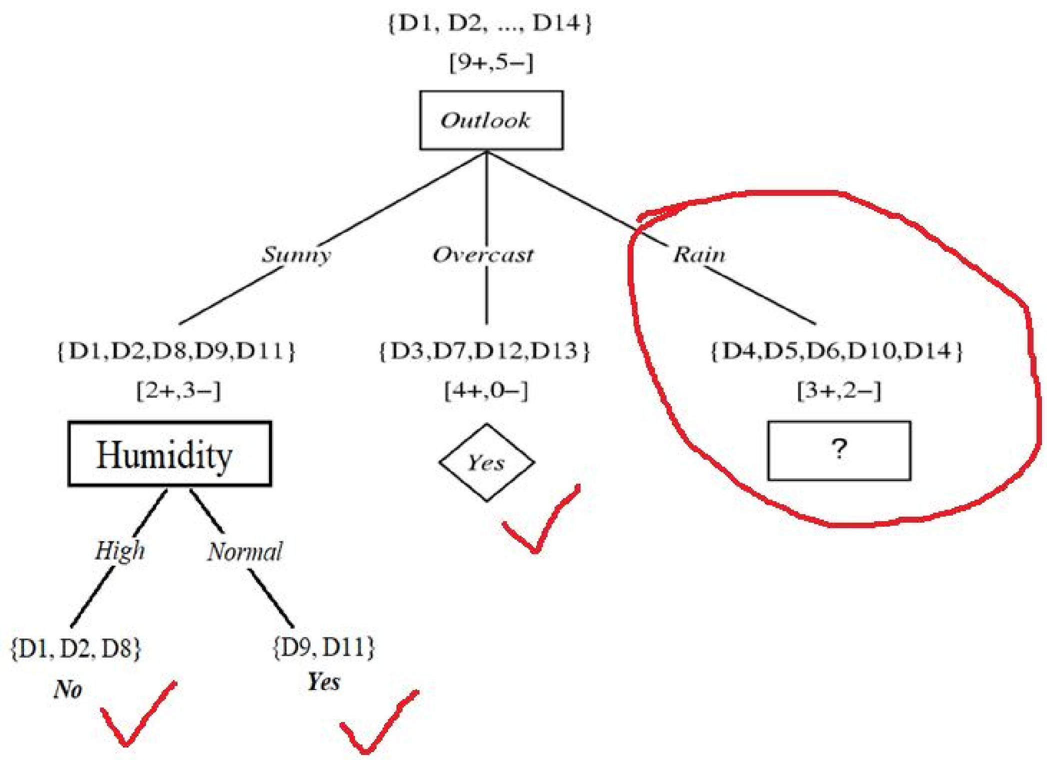


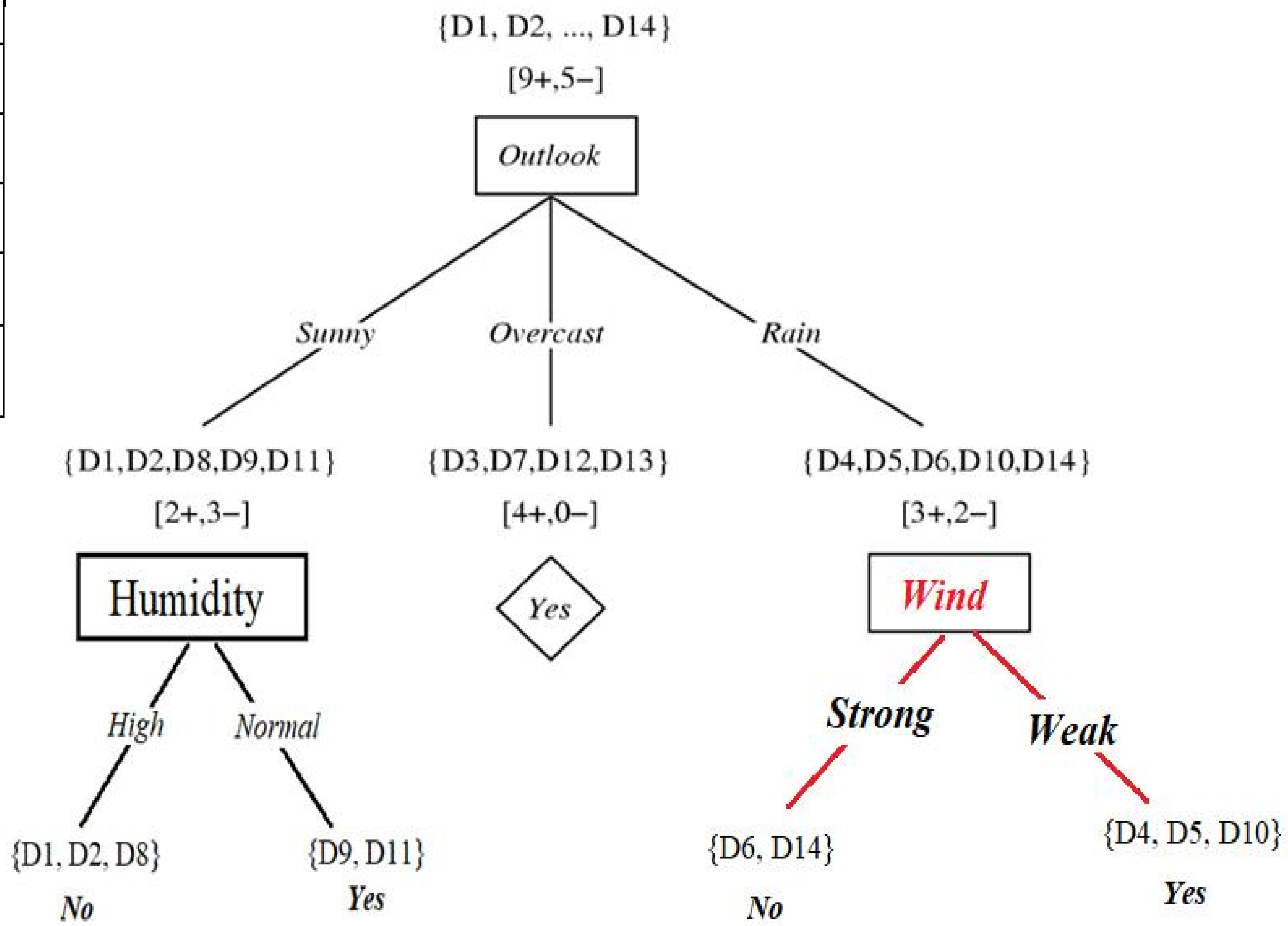
$Gain(S_{Rain}, Temperature) = 0.0192$

$Gain(S_{Rain}, Humidity) = 0.0192$

$\mathbf{Gain(S_{Rain}, Wind) = 0.97}$

Day	Temperature	Humidity	Wind	Play Tennis
D4	Mild	High	Weak	Yes
D5	Cool	Normal	Weak	Yes
D6	Cool	Normal	Strong	No
D10	Mild	Normal	Weak	Yes
D14	Mild	High	Strong	No





Homework

- Which feature will be at the root node of the decision tree trained for the following data? In other words which attribute makes a person most attractive?

Height	Hair	Eyes	Attractive?
small	blonde	brown	No
tall	dark	brown	No
tall	blonde	blue	Yes
tall	dark	Blue	No
small	dark	Blue	No
tall	red	Blue	Yes
tall	blonde	brown	No
small	blonde	blue	Yes

Decision Tree: Example

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