Part-1: Decision Tree Classifier

Decision Tree

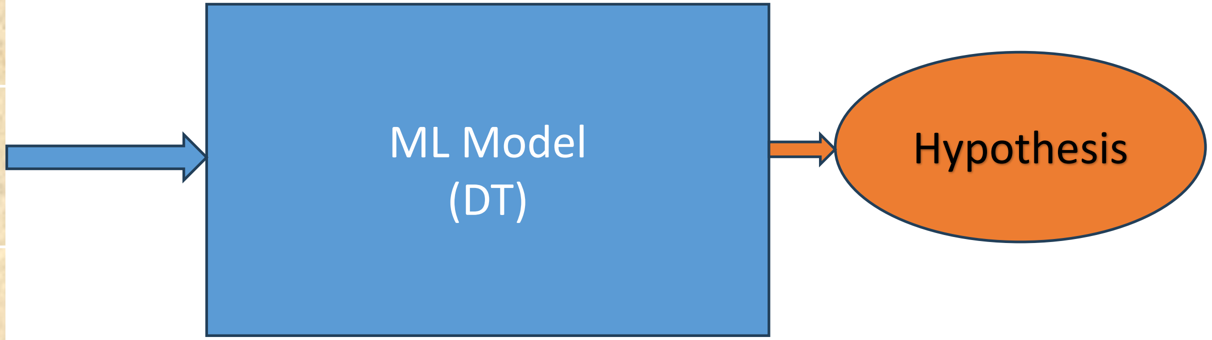
- Decision tree builds regression or classification models in the form of a tree structure.
  - For discrete value target functions (labels) – classification models
  - For continuous value target functions (labels) – regression models

- It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed.

- For discrete value target functions (for classification)
  - If-then-else rule
  - Most widely used approach

Recommending Mobile Applications

Gender	Age	Application
F	15	Pokemon go
F	25	WhatsApp
M	32	Chess
F	40	WhatsApp
M	12	Pokemon go
M	14	Pokemon go



Gender	Age	Application
F	18	?

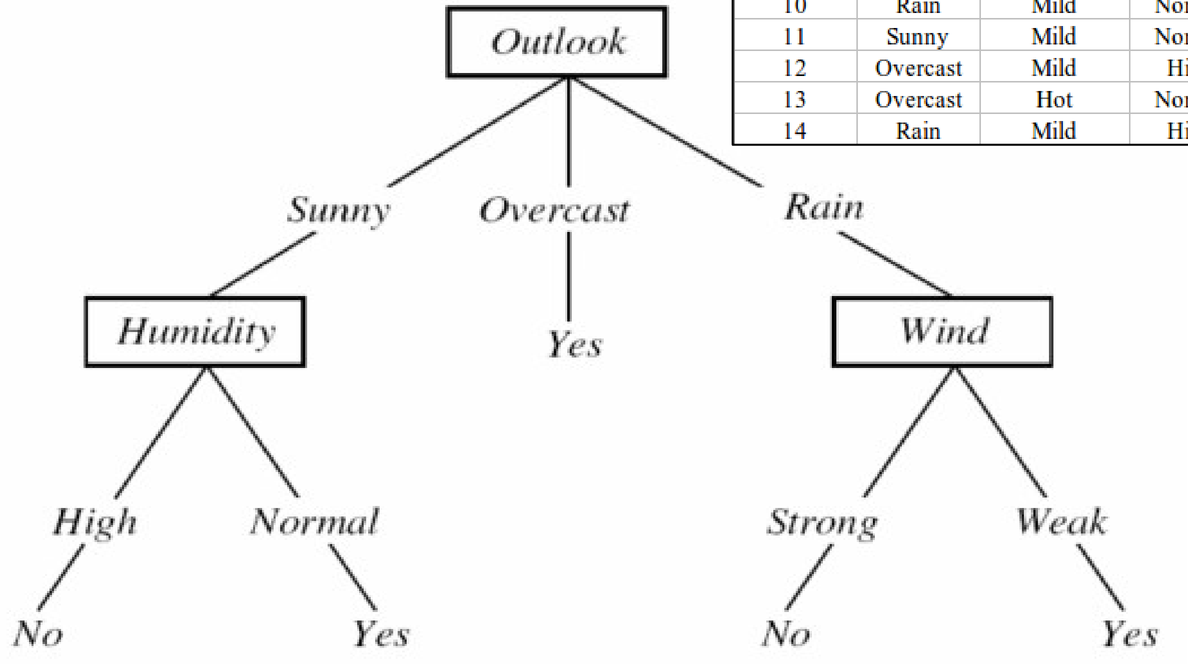
Decision Tree

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

Instance classification starts at the root node of the tree

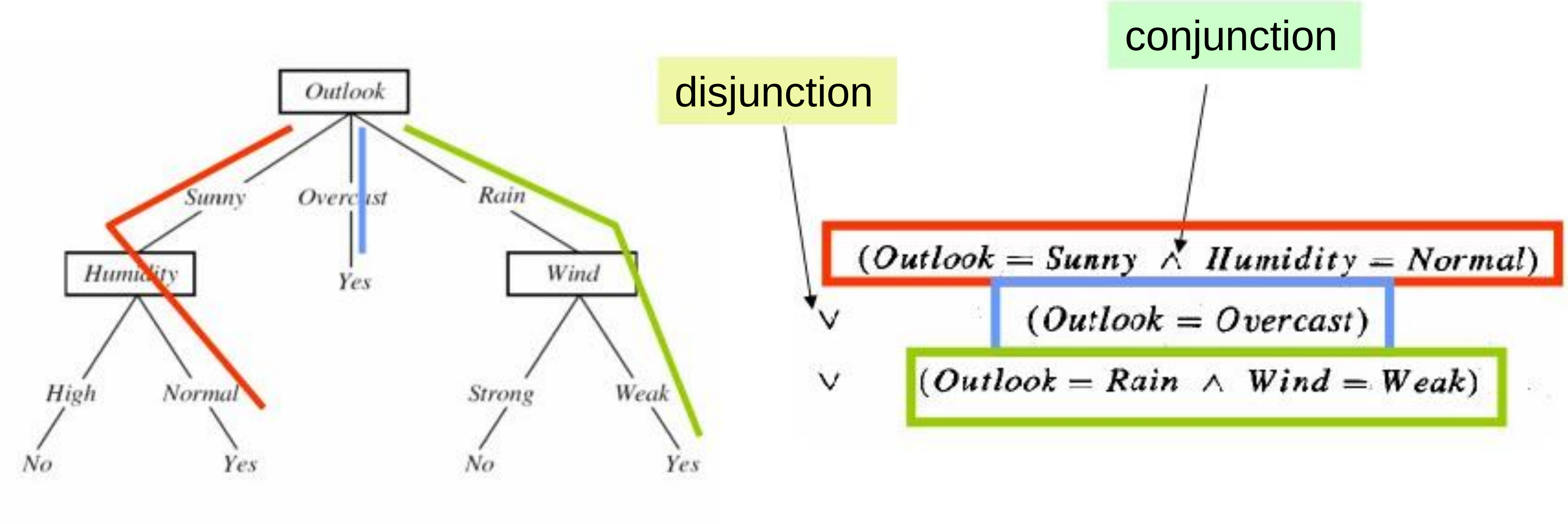
Testing the attribute values specified by this node

Moving down the tree branch corresponding to the value of the attribute in the given example.



This process is then repeated for the subtree rooted at the new node

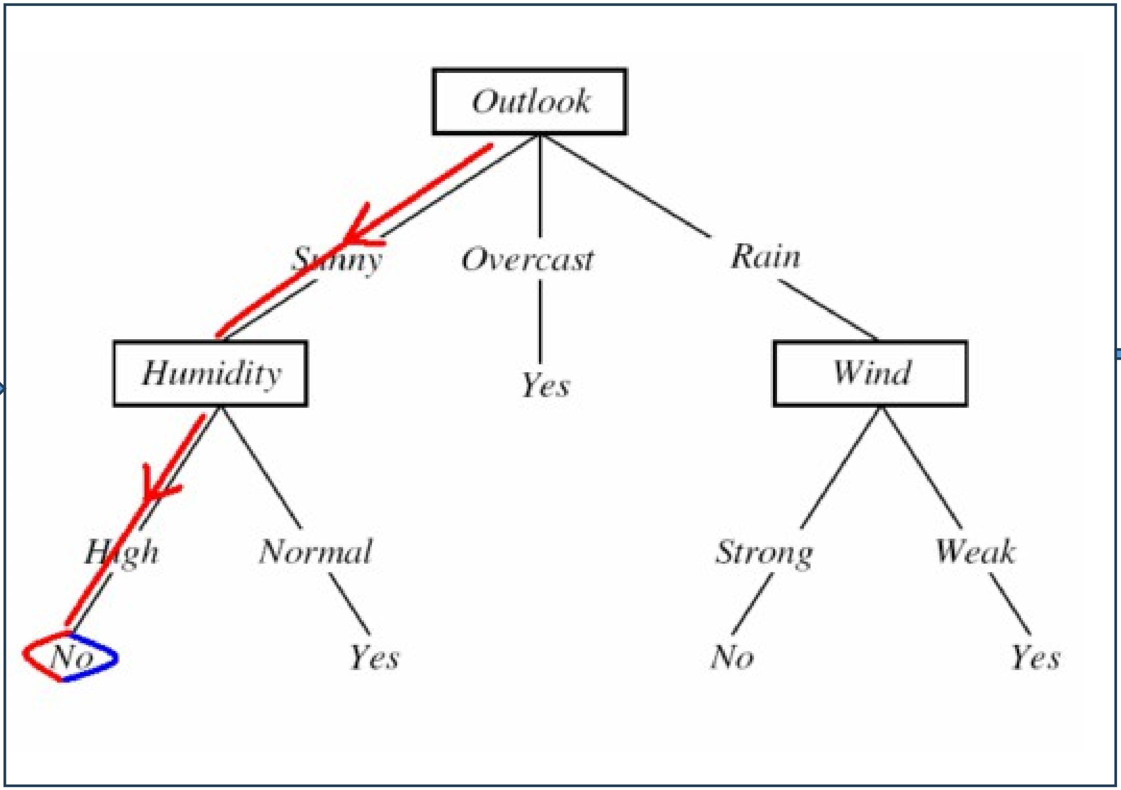
Decision Tree



- **Decision trees** represent a disjunction of conjunctions of constraints on the attribute values of instances.
- Each path from the tree root to a leaf corresponds to a conjunction of attribute tests, and the tree itself to a disjunction of these conjunctions.

Decision Tree

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



$$(Outlook = Sunny \wedge Humidity = Normal)$$
$$\vee \ (Outlook = Overcast)$$
$$\vee \ (Outlook = Rain \wedge Wind = Weak)$$

<Outlook = Sunny, Temperature = Hot, Humidity = High, Wind = Strong>

PlayTennis = No

Decision Tree

(a) $A \wedge \neg B$
(b) $A \vee [B \wedge C]$
(c) $A \ XOR \ B$
(d) $[A \wedge B] \vee [C \wedge D]$

Decision Trees (F = A ^ B')

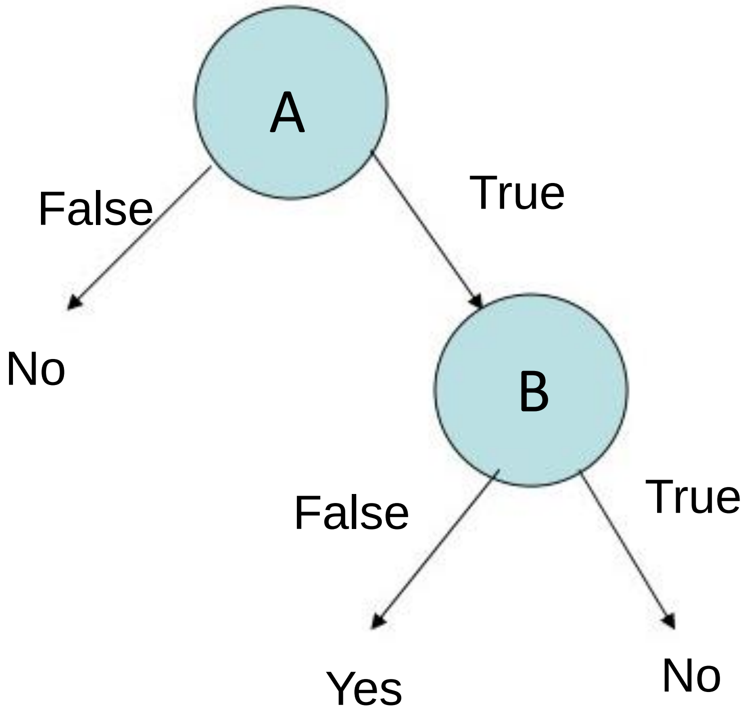
If (A=True and B = False) then Yes

else
    No

Decision Tree – assignment for IA

(a) $A \wedge \neg B$
(b) $A \vee [B \wedge C]$
(c) $A \ XOR \ B$
(d) $[A \wedge B] \vee [C \wedge D]$

Decision Trees (F = A ^ B')



If (A=True and B = False) then Yes

else
    No

Types of Problems Suitable for Decision Tree Classifier

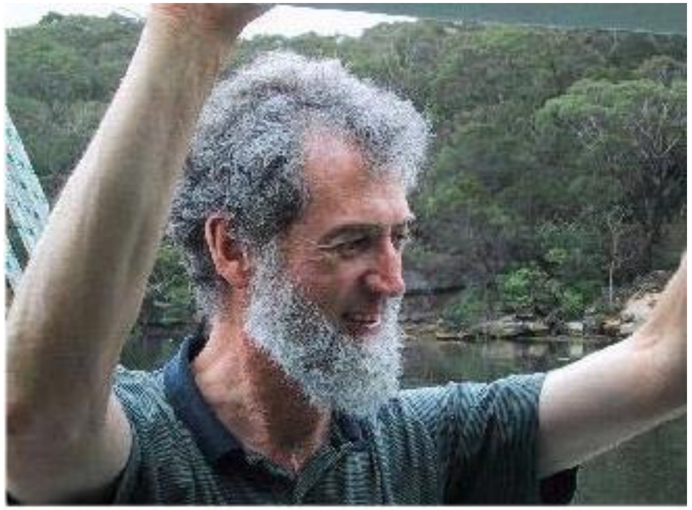
- *Instances are represented by multiple attribute values*
  - *Temperature: (Hot, Mild, Cold)*
- *The target function has discrete output values*
- *Disjunctive descriptions may be required*
- *The training data may contain errors*
- *The training data may contain missing attribute values*

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

Decision Tree Learning

ID3 stands for Iterative Dichotomiser 3

- This algorithm iteratively (repeatedly) dichotomizes(divides) features into two or more groups at each step.



Invented by Ross Quinlan

ID3 Algorithm (Quinlan 1986) and it's successors C4.5 and C5.0

- ***Employs a top-down***

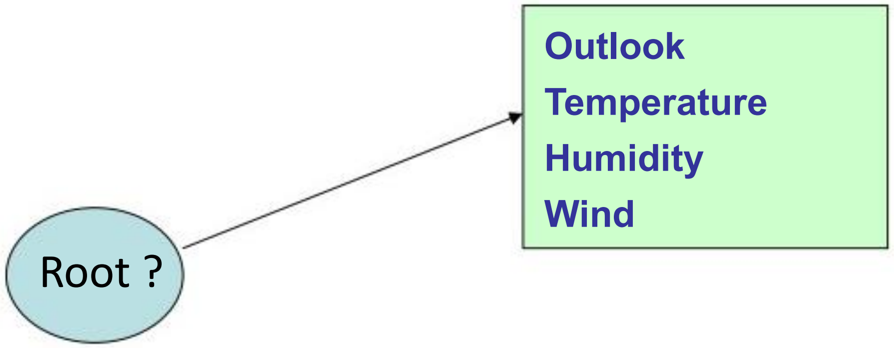
- An instance is classified by starting at the root node of the tree, testing the attribute specified by this node, then moving down the tree branch corresponding to the value of the attribute in the given example.

- This process is then repeated for the subtree rooted at the new node.

- Greedy search the space of possible decision trees.

- The algorithm never backtracks to reconsider earlier choices.

Which Attribute to Select ??



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

- Select the attribute that is **most useful** for classifying examples.
- What is a good quantitative measure of the worth of an attribute?

ID3 algorithm selecting a best attribute that yields **maximum Information Gain or minimum Entropy.**

ID3 Algorithm (Quinlan 1986)

- ID3 uses - **information gain** measure to select among the candidate attributes at each step while growing the tree.

- **Information gain** is based on information theory concept called ***Entropy***

- **Entropy** is a measure of the **uncertainty** associated with a random variable.

ID3 Algorithm

1. Calculate entropy for dataset.

2. For each attribute/feature.
   2.1. Calculate entropy for all its categorical values.
   2.2. Calculate information gain for the feature.

3. Find the feature with maximum information gain.

4. Repeat it until we get the desired tree.

- **Entropy in our Context**

  - Given a collection **S**, containing **positive and negative examples** of some target concept, the entropy of S relative to this boolean classification (**yes/ no**) is

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

  - Where $p_{\oplus}$ is the proportion of positive examples in **S** and $p_{\ominus}$ is the proportion of negative examples in S.

  - In all calculations involving entropy we define **0 log 0** to be **0**.

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

Entropy

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

- There are 14 examples. **9 positive** and **5 negative** examples [9+, 5-].
- The entropy of S relative to this Boolean (yes/no) classification is

$$
\begin{aligned}
Entropy([9+, 5-]) &= -(9/14)\log_2(9/14) - (5/14)\log_2(5/14) \\
&= 0.940
\end{aligned}
$$

Information Gain Measure

Information gain, is simply the expected reduction in entropy caused by partitioning the examples according to this attribute.

More precisely, the information gain, ***Gain(S, A)*** of ***an*** attribute **A,** relative to a collection of examples ***S,*** is defined as

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

where ***Values(A)*** is the set of all possible values for attribute **A**, and ***Sv,*** is the subset of **S** for which attribute **A** has value **v, i.e.,**

$$S_v = \{s \in S | A(s) = v\}$$

Gain (S, outlook)

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

Gain (S, Temperature)

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

Gain (S, Humidity)

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

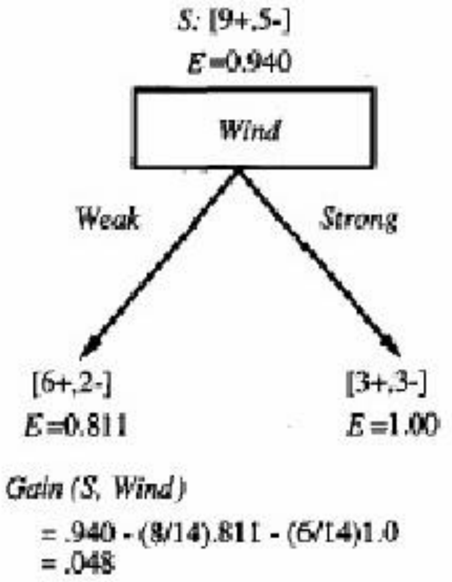
Gain (S, Wind)

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

Gain (S, Attribute = Wind)

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

$$
\begin{aligned}
Values(Wind) &= Weak, Strong \\
S &= [9+, 5-] \\
S_{Weak} &\leftarrow [6+, 2-] \\
S_{Strong} &\leftarrow [3+, 3-] \\
Gain(S, Wind) &= Entropy(S) - \sum_{v \in \{Weak, Strong\}} \frac{|S_v|}{|S|} Entropy(S_v) \\
&= Entropy(S) - (8/14) Entropy(S_{Weak}) \\
&\quad - (6/14) Entropy(S_{Strong}) \\
&= 0.940 - (8/14) 0.811 - (6/14) 1.00 \\
&= 0.048
\end{aligned}
$$



Gain (S,A)

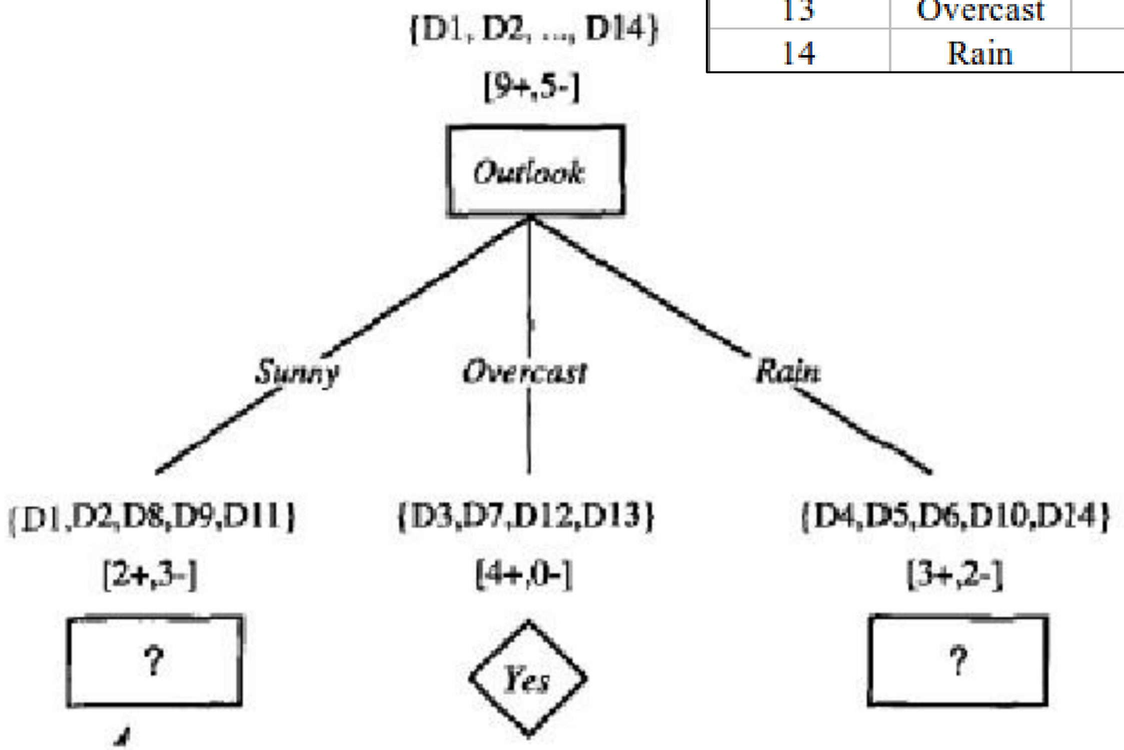
$$Gain(S, Outlook) = 0.246$$

$$Gain(S, Humidity) = 0.151$$

$$Gain(S, Wind) = 0.048$$

$$Gain(S, Temperature) = 0.029$$

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



Gain (S, Outlook) → Higher

Outlook is selected as root

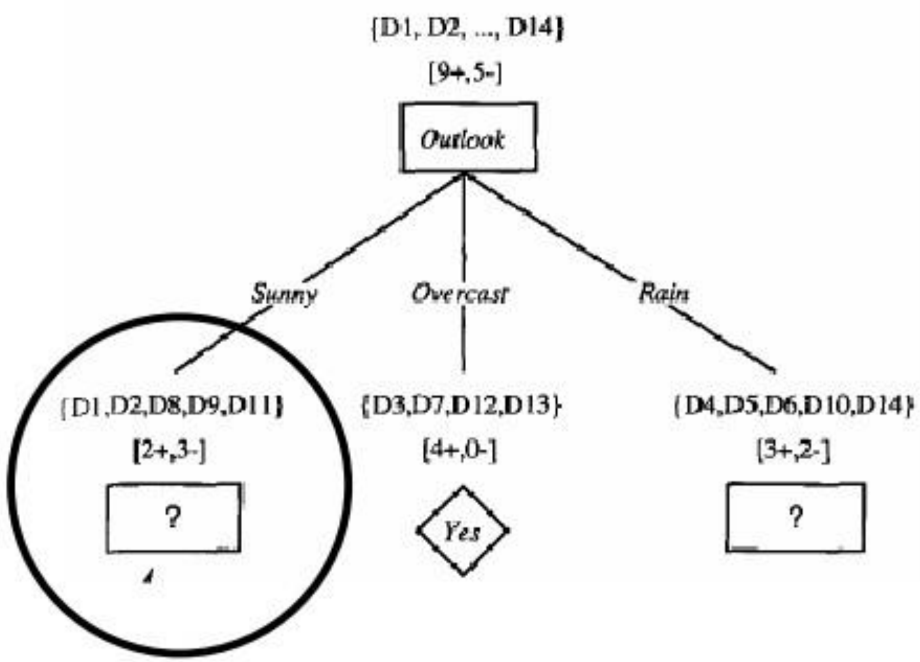
Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes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

Gain (S_{unny}, A)

A = Temperature, Humidity, Wind

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~~

Gain (Sunny, Temperature)



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

Temperature

(Hot) {0+, 2-)

(Mild) {1+, 1-}

(Cool) {1+, 0-}

Gain (S_{unny}, Temperature)

Entropy(S_{Sunny}) = - { 2/5 log(2/5) + 3/5 log(3/5)} = 0.97095

Temperature
(Hot) {0+, 2-)
(Mild) {1+, 1-}
(Cool) {1+, 0-}

Entropy(Hot) = 0

Entropy(Mild) = 1

Entropy(Cool) = 0

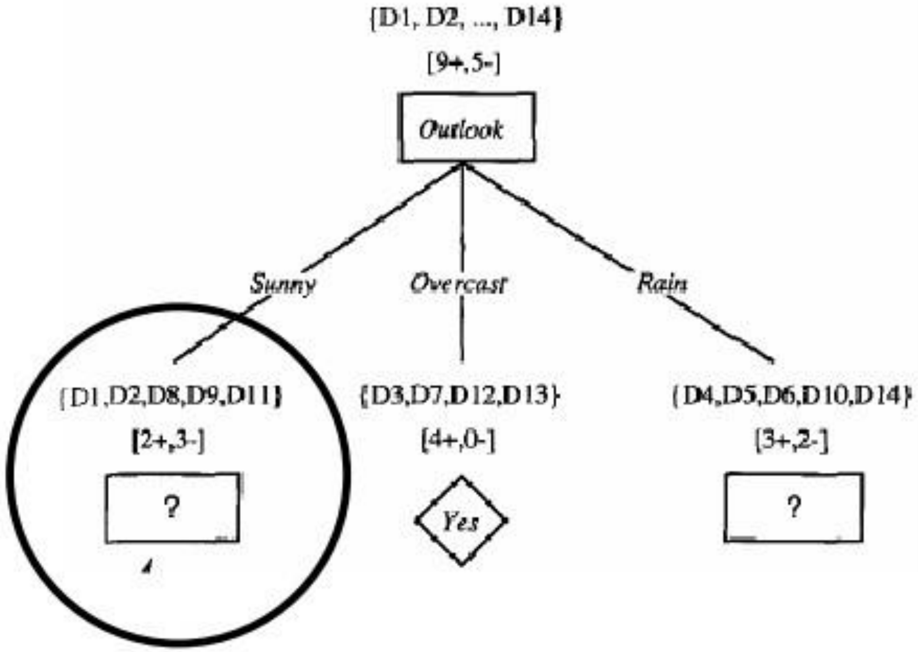
Gain(Su, Temperature) = 0.97095 - 2/5*0 - 2/5*1 - 1/5*0 = 0.57095

Gain (S_{unny}, Humidity)

A = Temperature, Humidity, Wind

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~~

Gain (S_{unny}, Humidity)



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

Humidity

(High) {0+, 3-}

(Normal) {2+, 0-}

Gain (S_{unny}, Humidity)

Entropy(S_{unny}) = - { 2/5 log(2/5) + 3/5 log(3/5)} = 0.97095

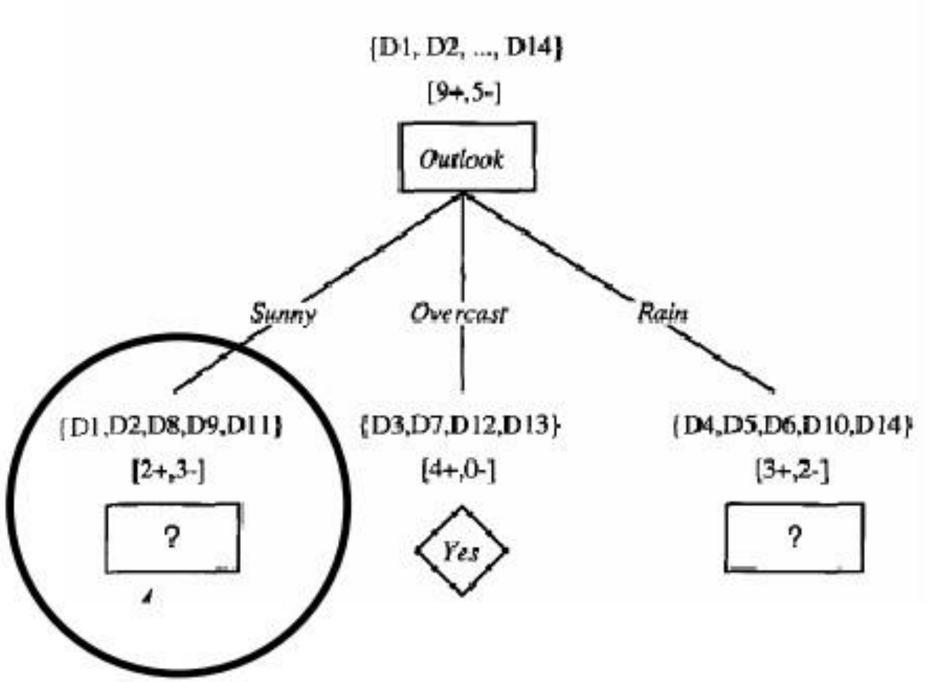
Humidity
(High) {0+, 3-}
(Normal) {2+, 0-}

Entropy(High) = 0

Entropy(Normal) = 0

Gain(Su, Humidity) = 0.97095 - 3/5*0 - 2/5*0 = **0.97095**

Gain (S_{unny}, Wind)



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~~

Temperature

(Hot) {0+, 2-)

(Mild) {1+, 1-}

(Cool) {1+, 0-}

Humidity

(High) {0+, 3-}

(Normal) {2+, 0-}

Wind

(Weak) {1+, 2-}

(Strong) {1+, 1-}

Gain (S_{unny},A)

Entropy(S_{unny}) = - { 2/5 log(2/5) + 3/5 log(3/5)} = 0.97095

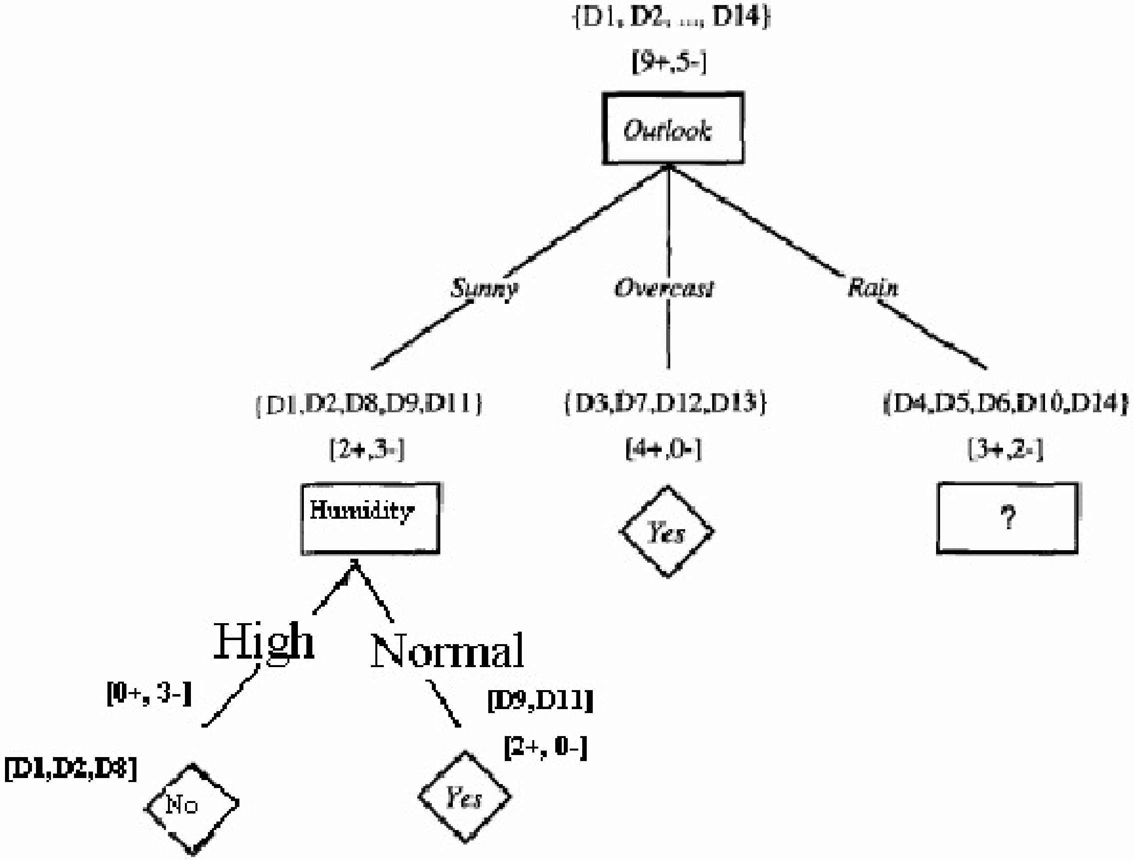
Wind
(Weak) {1+, 2-}
(Strong) {1+, 1-}

Entropy(Weak) = 0.9183

Entropy(Strong) = 1.0

Gain(S1, Wind) = 0.97095 - 3/5*0.9183 - 2/5*1 = 0.01997

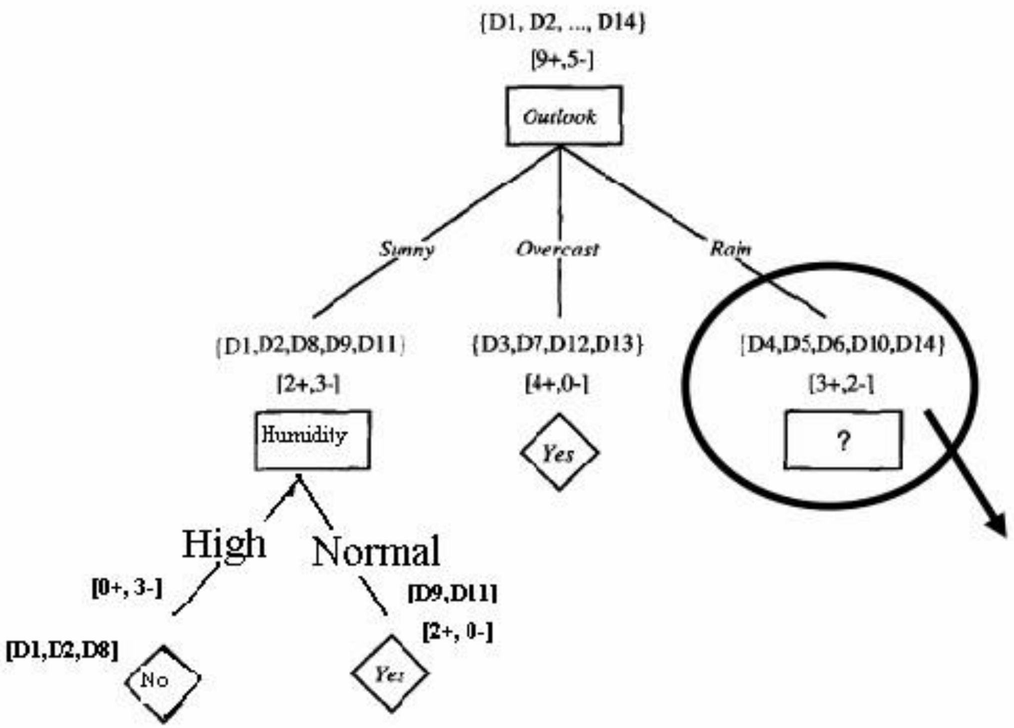
Modified Decision Tree



Gain (Rain, A)

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

Gain (S_{Rain},A)



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

Temperature

(Mild) {2+, 1-}

(Cool) {1+, 1-}

Humidity

(High) {1+, 1-}

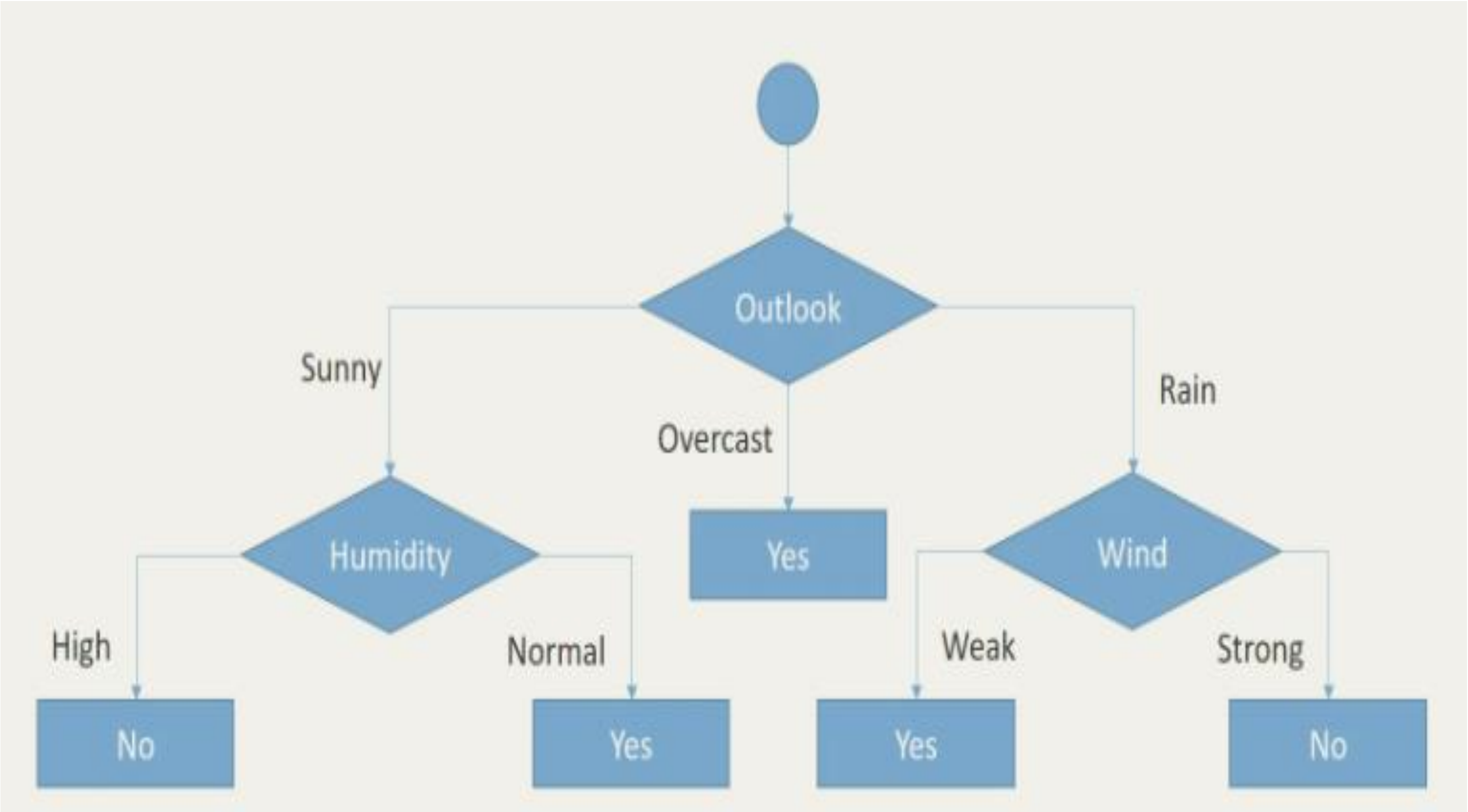
(Normal) {2+, 1-}

Wind

(Weak) {3+, 0-}

(Strong) {0+, 2-}

Final Decision Tree



Whether Jon will play tennis or not on D15?

	OL	Temp	Hum	Wind
D15	Rain	Mild	High	Weak ?

Assignment:

Obtain DT for the following data

Instance	Attribute_1	Attribute_2	Class
1	T	T	+
2	T	T	+
3	T	F	-
4	F	F	+
5	F	T	-
6	F	T	-

Capabilities and Limitations of ID3 Algorithm

Advantage:

- ID3 algorithm searches complete hypothesis space
- ID3 uses all training examples at each step in the search.
- The resulting search is much less sensitive to errors in individual training examples.

Capabilities and Limitations of ID3 Algorithm

- Disadvantage:
  - ID3 maintain only a single current hypothesis as it searches through the space of decision trees. Loses the capabilities of representing all consistent hypothesis.
  - ID3 performs **no backtracking** in its search.
    - Once it selects an attribute to test at a particular level in the tree, it never backtracks to reconsider this choice.
    - Therefore, converging to locally optimal solutions that are not globally optimal.

Thank you