

ID3 Algorithm

EECS 349 Machine Learning
Recitation

Bongjun Kim

Training examples

<i>Day</i>	<i>Outlook</i>	<i>Temperature</i>	<i>Humidity</i>	<i>Wind</i>	<i>PlayTennis</i>
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

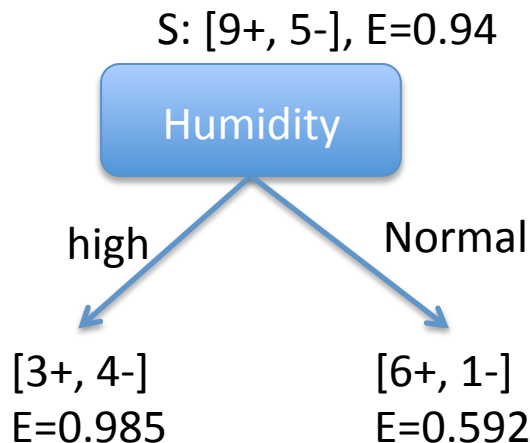
How to choose the best attribute

- Information Gain

- The expected reduction in Entropy

- Entropy (before split) – Entropy (after split)

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

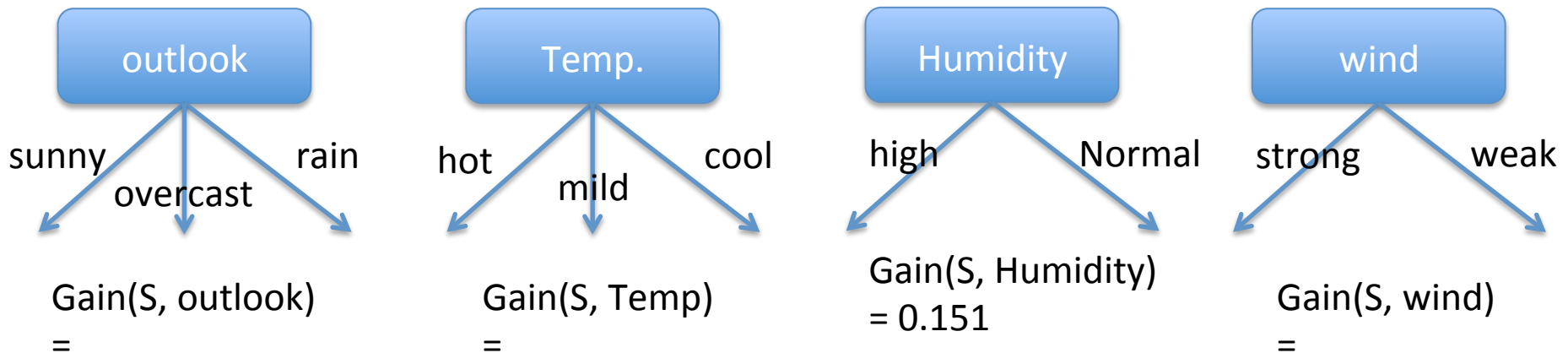


$$\begin{aligned} Gain(S, Humidity) &= 0.94 - ((7/14)*0.985 + (7/14)*0.592) \\ &= 0.151 \end{aligned}$$

How to choose the best attribute

- Information Gain
 - The expected reduction in Entropy
 - Entropy (before split) – Entropy (after split)

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



Entropy

- Entropy

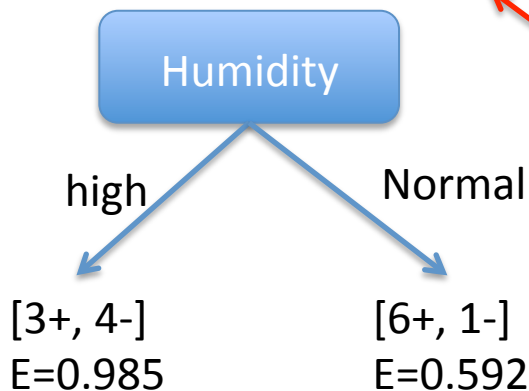
- P_1 = probability I will play tennis

- P_2 = probability I will NOT play tennis

$$Entropy(S) = \sum_j -P_j \log_2 P_j$$

$$= -P_1 \log_2 P_1 - P_2 \log_2 P_2$$

S: [9+, 5-], E=0.94



Entropy([9+, 5-])

$$= -(9/14)\log_2(9/14) - (5/14)\log_2(5/14)$$

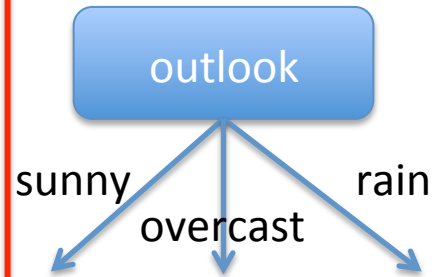
$$= 0.94$$

Let's build DT

<i>Day</i>	<i>Outlook</i>	<i>Temperature</i>	<i>Humidity</i>	<i>Wind</i>	<i>PlayTennis</i>
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

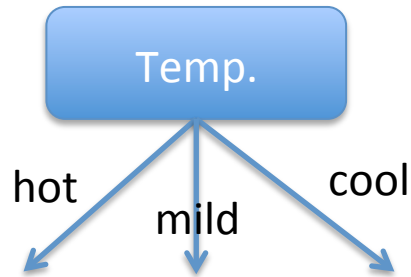
Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal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

S: [9+, 5-], E=0.94



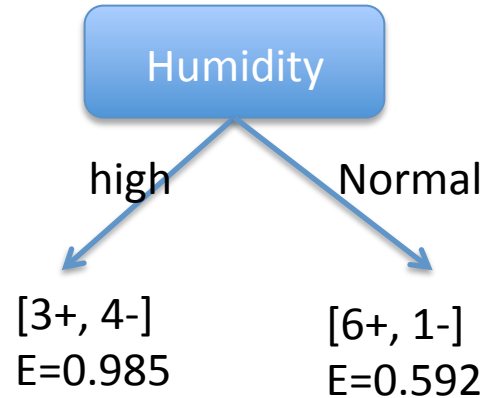
Gain(S, outlook)
= 0.246

S: [9+, 5-], E=0.94



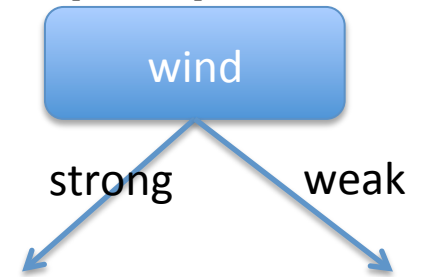
Gain(S, Temp)
= 0.029

S: [9+, 5-], E=0.94



Gain(S, Humidity)
= 0.151

S: [9+, 5-], E=0.94



Gain(S, wind)
= 0.048

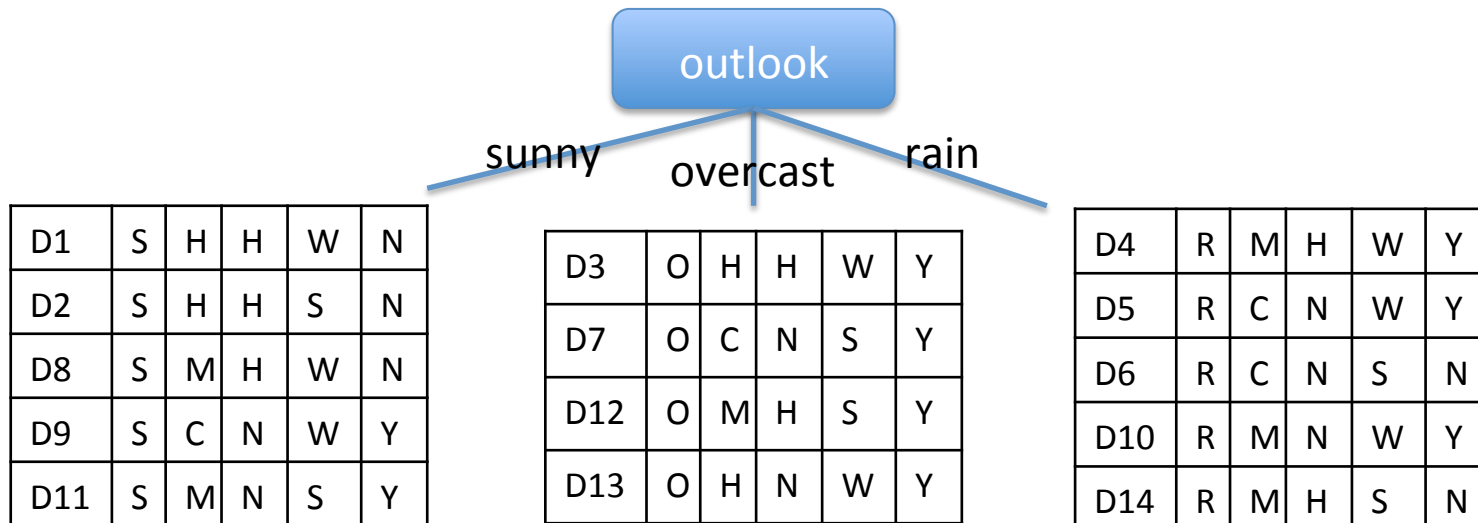
Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	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) = 0.246

Gain(S, Temperature)=0.029

Gain(S, humidity) = 0.151

Gain(S, wind) = 0.048



outlook temperature humidity wind

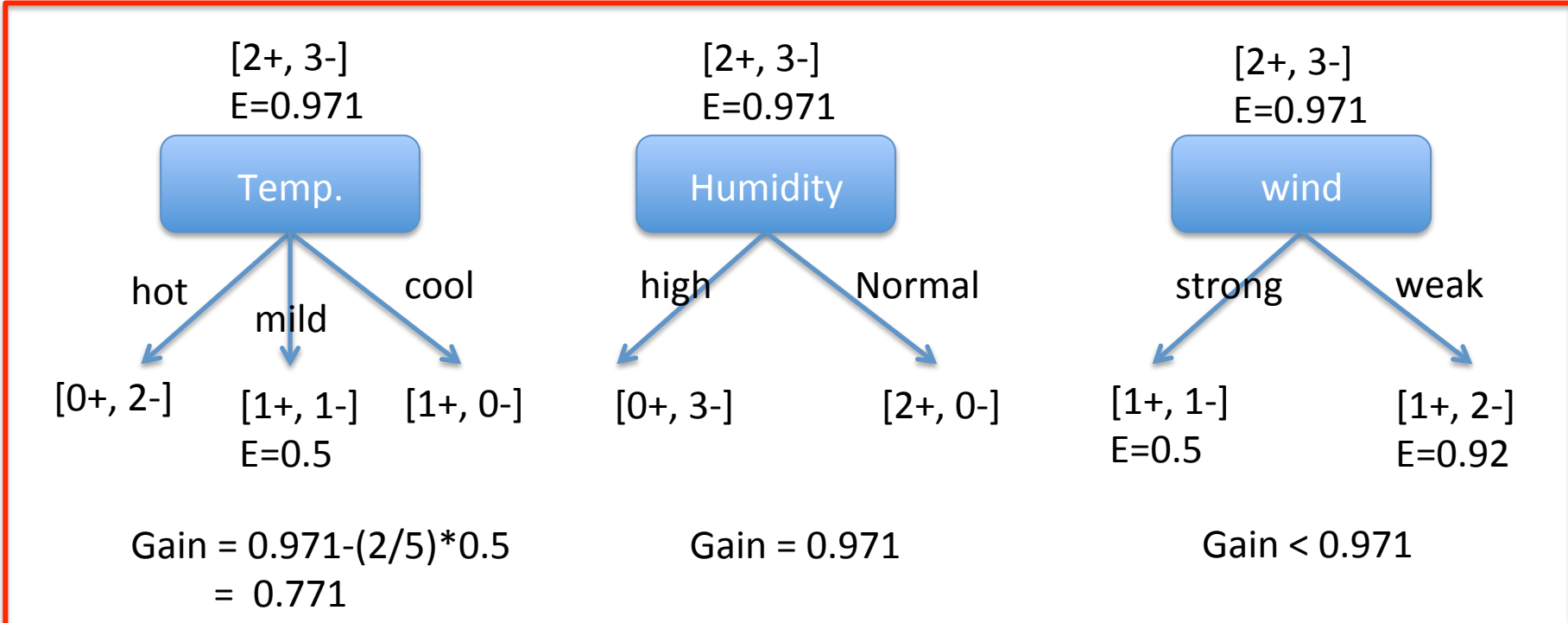
outlook

sunny overcast rain

D1	S	H	H	W	N
D2	S	H	H	S	N
D8	S	M	H	W	N
D9	S	C	N	W	Y
D11	S	M	N	S	Y

D3	O	H	H	W	Y
D7	O	C	N	S	Y
D12	O	M	H	S	Y
D13	O	H	N	W	Y

D4	R	M	H	W	Y
D5	R	C	N	W	Y
D6	R	C	N	S	N
D10	R	M	N	W	Y
D14	R	M	H	S	N



outlook temperature humidity wind

outlook

sunny overcast rain

D1	S	H	H	W	N
D2	S	H	H	S	N
D8	S	M	H	W	N
D9	S	C	N	W	Y
D11	S	M	N	S	Y

D3	O	H	H	W	Y
D7	O	C	N	S	Y
D12	O	M	H	S	Y
D13	O	H	N	W	Y

D4	R	M	H	W	Y
D5	R	C	N	W	Y
D6	R	C	N	S	N
D10	R	M	N	W	Y
D14	R	M	H	S	N

Humidity

high Normal

D1	S	H	H	W	N
D2	S	H	H	S	N
D8	S	M	H	W	N

No

D9	S	C	N	W	Y
D11	S	M	N	S	Y

Yes

outlook temperature humidity wind

outlook

sunny overcast rain

D1	S	H	H	W	N
D2	S	H	H	S	N
D8	S	M	H	W	N
D9	S	C	N	W	Y
D11	S	M	N	S	Y

D3	O	H	H	W	Y
D7	O	C	N	S	Y
D12	O	M	H	S	Y
D13	O	H	N	W	Y

D4	R	M	H	W	Y
D5	R	C	N	W	Y
D6	R	C	N	S	N
D10	R	M	N	W	Y
D14	R	M	H	S	N

Humidity

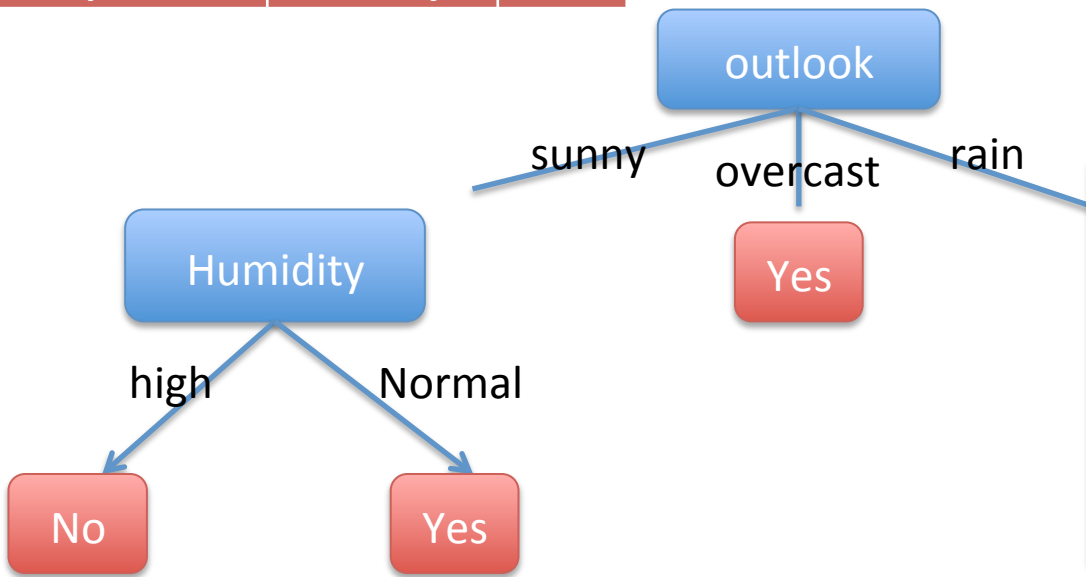
high Normal

No

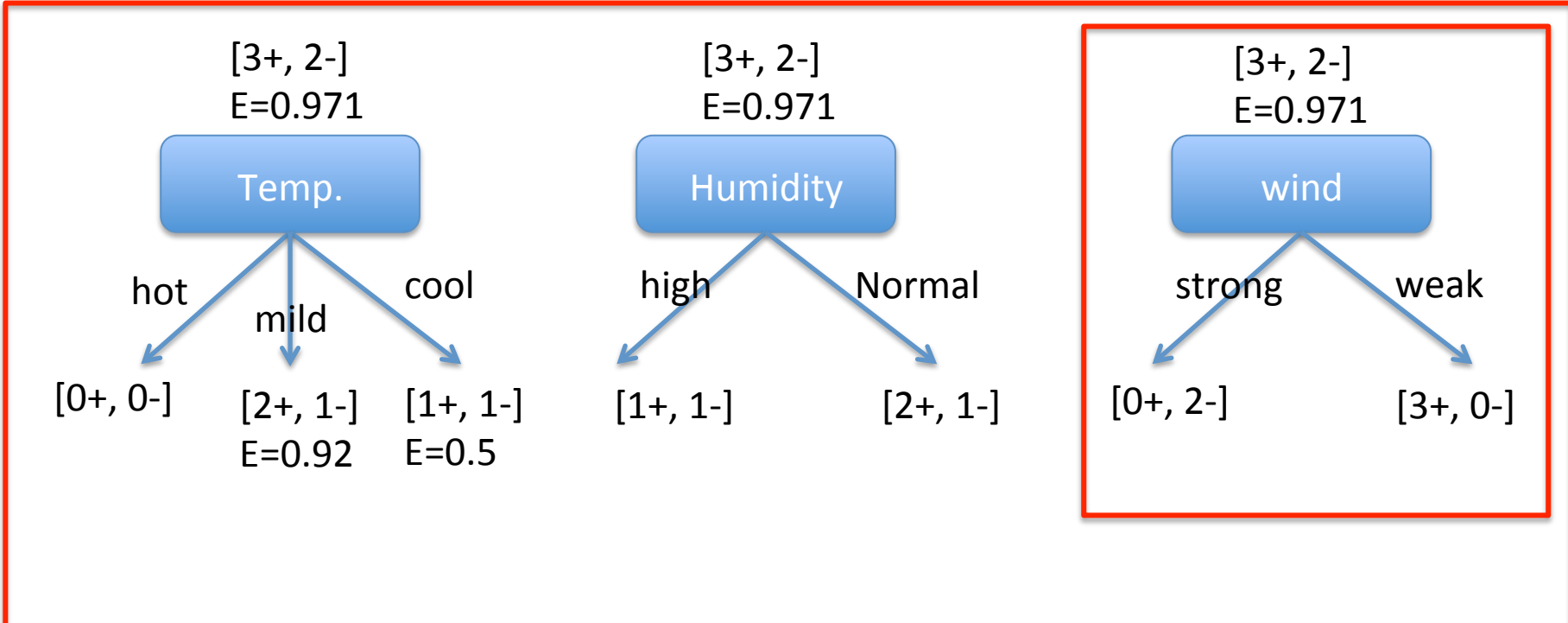
Yes

Yes

outlook temperature humidity wind



D4	R	M	H	W	Y
D5	R	C	N	W	Y
D6	R	C	N	S	N
D10	R	M	N	W	Y
D14	R	M	H	S	N



outlook temperature humidity wind

outlook

sunny

overcast

rain

D1	S	H	H	W	N
D2	S	H	H	S	N
D8	S	M	H	W	N
D9	S	C	N	W	Y
D11	S	M	N	S	Y

D3	O	H	H	W	Y
D7	O	C	N	S	Y
D12	O	M	H	S	Y
D13	O	H	N	W	Y

D4	R	M	H	W	Y
D5	R	C	N	W	Y
D6	R	C	N	S	N
D10	R	M	N	W	Y
D14	R	M	H	S	N

Humidity

high

Normal

No

Yes

Yes

wind

strong

weak

D6	R	C	N	S	N
D14	R	M	H	S	N

D4	R	M	H	W	Y
D5	R	C	N	W	Y
D10	R	M	N	W	Y

outlook temperature humidity wind

outlook

sunny overcast rain

D1	S	H	H	W	N
D2	S	H	H	S	N
D8	S	M	H	W	N
D9	S	C	N	W	Y
D11	S	M	N	S	Y

D3	O	H	H	W	Y
D7	O	C	N	S	Y
D12	O	M	H	S	Y
D13	O	H	N	W	Y

D4	R	M	H	W	Y
D5	R	C	N	W	Y
D6	R	C	N	S	N
D10	R	M	N	W	Y
D14	R	M	H	S	N

Humidity

high Normal

No

Yes

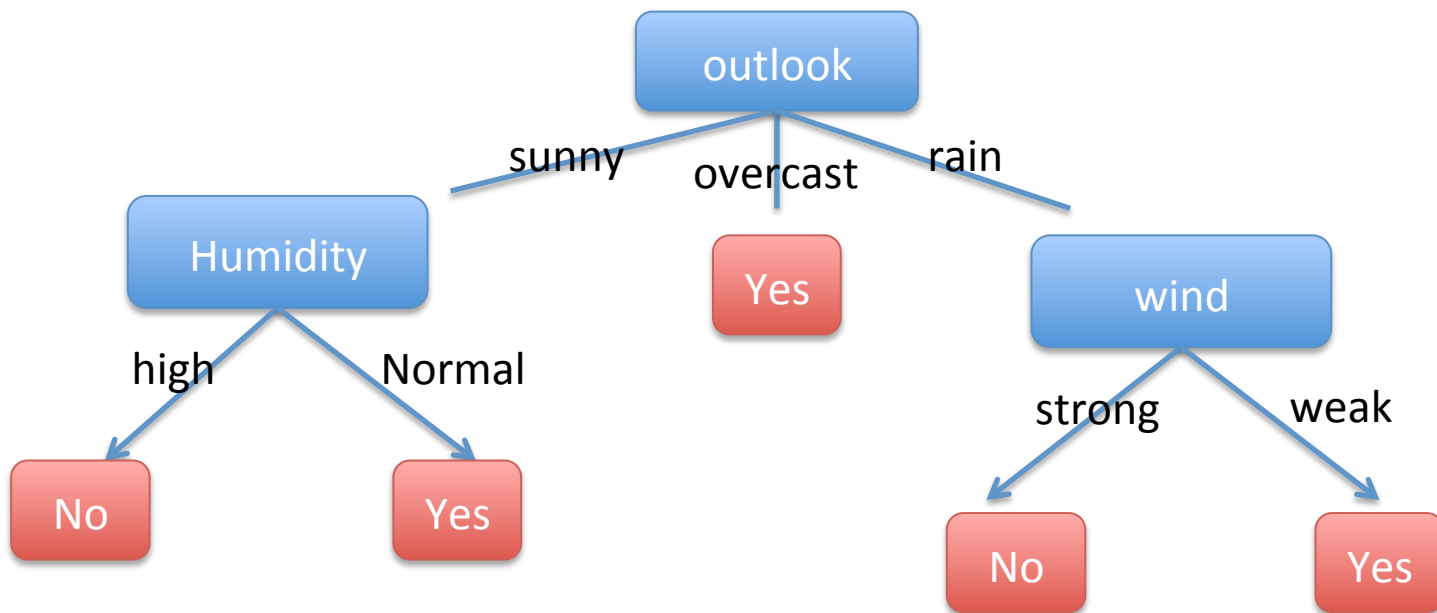
Yes

wind

strong weak

No

Yes



outlook	temperature	humidity	wind	play
rainy	Mild	High	weak	?