

Outlook	Temp	Humidity	Wind	Playing
1 Sunny	Hot	High	False	no
2 Sunny	Hot	High	True	no
3 Overcast	Hot	High	F	yes
4 Raining	Mild	High	F	yes
5 Raining	Cool	Normal	F	yes
6 Raining	Cool	Normal	T	no
7 Overcast	Cool	Normal	T	yes
8 Sunny	Mild	High	F	no
9 Sunny	Cool	Normal	F	yes
10 Raining	Mild	Normal	F	yes
11 Sunny	Mild	Normal	T	yes
12 Overcast	Mild	High	T	yes
13 Overcast	Hot	Normal	F	yes
14 Raining	Mild	High	F	no

$$E(\text{outlook}) = -\left(\frac{9}{14}\right) \log\left(\frac{9}{14}\right) - \left(\frac{5}{14}\right) \log\left(\frac{5}{14}\right)$$

$$\text{Gini}(S) = 1 - \left[\left(\frac{9}{14} \right)^2 + \left(\frac{5}{14} \right)^2 \right] = 0.4591$$

		yes	no	Total
outlook	Sunny	2	3	5
	Overcast	4	0	4
	Rainy	2	3	5
				14

		yes	no	Total
Temp	Hot	2	2	4
	Mild	4	2	6
	Cool	3	1	4
				14

$$\begin{aligned} \text{Gini}(S, \text{outlook}) &= \frac{5}{14} \text{gini}(2,3) + \frac{4}{14} \text{gini}(4,0) + \frac{5}{14} \text{gini}(2,3) \\ &= \left(\frac{5}{14} \left[1 - \left(\frac{2}{5} \right)^2 - \left(\frac{3}{5} \right)^2 \right] \right) + \left(\frac{4}{14} \times 0 \right) + \left(\frac{5}{14} \left[1 - \left(\frac{2}{5} \right)^2 - \left(\frac{3}{5} \right)^2 \right] \right) \\ &= 0.342 \end{aligned}$$

$$\text{Gini gain}(S, \text{outlook}) = 0.459 - 0.342 = \boxed{0.117}$$

$$\begin{aligned} \text{Gini}(S, \text{temp}) &= \frac{4}{14} \text{gini}(2,2) + \frac{6}{14} \text{gini}(4,2) + \frac{4}{14} \text{gini}(3,1) \\ &= \frac{4}{14} \left(1 - \left(\frac{2}{4} \right)^2 - \left(\frac{2}{4} \right)^2 \right) + \frac{6}{14} \left(1 - \left(\frac{4}{6} \right)^2 - \left(\frac{2}{6} \right)^2 \right) + \frac{4}{14} \left(1 - \left(\frac{3}{4} \right)^2 - \left(\frac{1}{4} \right)^2 \right) \\ &= 0.443 \end{aligned}$$

$$\begin{aligned} \text{Gini gain}(S, \text{temp}) &= 0.459 - 0.443 \\ &= \boxed{0.016} \end{aligned}$$

$$\text{Gini gain}(S, \text{humidity}) = 0.459 - 0.3674 = \boxed{0.0916}$$

$$\text{Gini gain}(S, \text{wind}) = 0.459 - 0.4286 = \boxed{0.0304}$$

Select the feature having largest gini gain \Rightarrow outlook
or smallest gini index

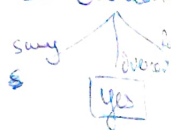
Under sunny

Outlook	Temp	Humidity	Wind	
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Sunny	Mild	High	Weak	No
Sun	Cool	Norm	Weak	yes
Sunny	Mild	Norm	Strong	yes

Hum	High	0	3	3
Norm	2	0	0	2

		yes	No	Tot
Temp	Hot	0	2	2
	Cool	1	1	2
	Mild	1	0	1
				5

Wind	Weak	1	2	3
Strong	1	1	2	



$$\begin{aligned} \text{gini}(\text{sunny}, \text{temp}) &= \frac{2}{5} \text{gini}(0,2) + \frac{2}{5} \text{gini}(1,1) + \frac{1}{5} \text{gini}(1,0) \\ &= \frac{2}{5} \left[1 - \left(\frac{0}{2}\right)^2 - \left(\frac{2}{2}\right)^2 \right] + \frac{2}{5} \left[1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 \right] + \frac{1}{5} \times 0 \\ &= \frac{2}{5} [1 - 0.25 - 0.25] = 0.2 \end{aligned}$$

$$\text{gini}(\text{sunny}, \text{hum}) = \frac{3}{5} \text{gini}(0,3) + \frac{2}{5} \text{gini}(2,0) = 0$$

$$\begin{aligned} \text{gini}(\text{sunny}, \text{wind}) &= \frac{3}{5} \text{gini}(1,2) + \frac{2}{5} \text{gini}(1,1) \\ &= \frac{3}{5} \left[1 - \left(\frac{1}{3}\right)^2 - \left(\frac{2}{3}\right)^2 \right] + \frac{2}{5} \left[1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 \right] \\ &= \frac{3}{5} [1 - 0.11 - 0.44] + 0.2 = 0.27 + 0.2 = 0.47 \end{aligned}$$

Humidity has lower gini index

Under rainy

		yes	No	Tot
Temp	Hot	2	1	3
	Cool	1	1	2
				5

		yes	No	Tot
Hum	High	1	1	2
	Norm	2	1	3
				5

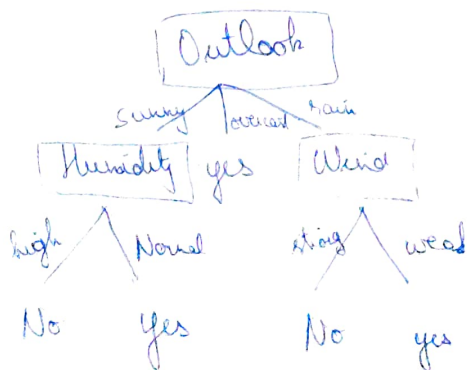
		yes	No	Tot
Wind	Strong	0	2	2
	Weak	3	0	3
				5

$$\begin{aligned} \text{gini}(\text{rainy}, \text{temp}) &= \frac{3}{5} \text{gini}(2,1) + \frac{2}{5} \text{gini}(1,1) \\ &= 0.47 \end{aligned}$$

$$\begin{aligned} \text{gini}(\text{rainy}, \text{hum}) &= \frac{2}{5} \text{gini}(1,1) + \frac{3}{5} \text{gini}(2,1) \\ &= \frac{2}{5} \left[1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 \right] + \frac{3}{5} \left[1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 \right] \\ &= 0.47 \end{aligned}$$

$$\begin{aligned} \text{gini}(\text{rainy}, \text{wind}) &= \frac{2}{5} \text{gini}(0,2) + \frac{3}{5} \text{gini}(3,0) \\ &= 0 \end{aligned}$$

Wind has lower index



$$\begin{aligned} \text{Gini}(\text{sunny}, \text{temp}) &= \frac{2}{5} \text{gini}(0,2) + \frac{2}{5} \text{gini}(1,1) + \frac{1}{5} \text{gini}(1,0) \\ &= \frac{2}{5} \left[1 - \left(\frac{0}{2}\right)^2 - \left(\frac{2}{2}\right)^2 \right] + \frac{2}{5} \left[1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 \right] + \frac{1}{5} \times 0 \\ &= \frac{2}{5} [1 - 0.25 - 0.25] = 0.2 \end{aligned}$$

$$\text{Gini}(\text{sunny}, \text{hum}) = \frac{3}{5} \text{gini}(0,3) + \frac{2}{5} \text{gini}(2,0) = 0$$

$$\begin{aligned} \text{Gini}(\text{sunny}, \text{wind}) &= \frac{3}{5} \text{gini}(1,2) + \frac{2}{5} \text{gini}(1,1) \\ &= \frac{3}{5} \left[1 - \left(\frac{1}{3}\right)^2 - \left(\frac{2}{3}\right)^2 \right] + \frac{2}{5} \left[1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 \right] \\ &= \frac{3}{5} [1 - 0.11 - 0.44] + 0.2 = 0.27 + 0.2 = 0.47 \end{aligned}$$

Humidity has lower gini index

Under rainy

Temp	Held	yes No Tot	
		2	1 3
	Cool	1	1 2
		5	

Hum	High	yes No Tot	
		1	1 2
	Norm	2	1 3
		5	

Wind	Strong	yes No Tot	
		0	2 2
	Weak	3	0 3
		5	

$$\begin{aligned} \text{Gini}(\text{rainy}, \text{temp}) &= \frac{3}{5} \text{gini}(2,1) + \frac{2}{5} \text{gini}(1,1) \\ &= 0.47 \end{aligned}$$

$$\begin{aligned} \text{Gini}(\text{rainy}, \text{hum}) &= \frac{2}{5} \text{gini}(1,1) + \frac{3}{5} \text{gini}(2,1) \\ &= \frac{2}{5} \left(1 - \left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^2 \right) + \frac{3}{5} \left(1 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2 \right) \\ &= 0.47 \end{aligned}$$

$$\begin{aligned} \text{gini}(\text{rainy}, \text{wind}) &= \frac{2}{5} \text{gini}(0,2) + \frac{3}{5} \text{gini}(3,0) \\ &= 0 \end{aligned}$$

Wind has lower index

