

Rizki Hanif K
 All. 2022.14644

No. _____

Date: _____

5001

Outlook	temperature	Humidity	windy	play
Sunny	Hot	high	No	Don't play
Sunny	Hot	high	Yes	Don't play
Cloudy	Hot	high	no	play
Rainy	Mild	high	no	play
Rainy	cool	normal	no	play
Rainy	cool	normal	Yes	play
cloudy	mild	normal	Yes	play
Sunny	cool	High	no	Don't play
Sunny	mild	normal	no	play
Rainy	mild	normal	no	play
Sunny	mild	normal	Yes	play
cloudy	mild	High	Yes	play
cloudy	Hot	Normal	no	play
Rainy	mild	High	Yes	Don't play

perhitungan Entropy total :

$$\text{Entropy}(\text{tot}) = \left(-\frac{4}{14} \times \log_2 \left(\frac{4}{14} \right) \right) + \left(-\frac{10}{14} \times \log_2 \left(\frac{10}{14} \right) \right)$$

$$= -0,2857 \cdot \log_2 0,2857 + (-0,7143) \cdot \log_2 0,7143$$

$$= \frac{-0,2857 \cdot -0,5440}{0,3010} + \frac{(-0,7143) \cdot -0,1461}{0,3010}$$

$$= -0,2857 \cdot -1,8073 + (-0,7143) \cdot -0,4853$$

$$= 0,5163 + 0,347$$

$$= 0,8631$$

Perhitungan node 1

Node1		dim/kasus	Don't play	play	Entropy	Gam
Total		14	4	10	0,8621	
Outlook						0,2588
	sunny	5	3	2	0,9707	
	cloudy	9	0	4	0	
	rainy	5	1	4	0,7217	
temp						1,838
	Hot	4	2	1		
	mild	6			0,9186	
	cool	4	0	4	0	
Humidity						0,3706
	high	7	4	3	0,985	
	normal	7	0	7	0	
windy						0,0061
	no	8	2	6	0,8111	
	yes	6	2	4	0,9186	

➤ Perhitungan entropy outlook

1. Sunny

$$\begin{aligned} \text{Entroph (Sunny)} &= \left(-\frac{3}{5} \times \log_2 \left(\frac{3}{5} \right) \right) + \left(-\frac{2}{5} \times \log_2 \left(\frac{2}{5} \right) \right) \\ &= -0.6 \log_2(0.6) + (-0.4) \times \log_2(0.4) \\ &= -0.6 \cdot -0.7368 + (-0.4) \cdot -1.3219 \\ &= 0.4420 + 0.5287 \\ &= 0.9707 \end{aligned}$$

2. Cloudy

$$\begin{aligned} \text{Entroph (cloudy)} &= \left(-\frac{0}{4} \times \log_2 \left(\frac{0}{4} \right) \right) + \left(-\frac{4}{4} \times \log_2 \left(\frac{4}{4} \right) \right) \\ &= 0 \cdot \log_2 0 + (-1) \cdot \log_2 1 \\ &= 0 \cdot 0 + (-1) \cdot 0 \\ &= 0 + 0 = 0 \end{aligned}$$

3. Rainy

$$\begin{aligned} \text{Entroph (Rainy)} &= \left(-\frac{1}{3} \times \log_2 \left(\frac{1}{3} \right) \right) + \left(-\frac{2}{3} \times \log_2 \left(\frac{2}{3} \right) \right) \\ &= -0.2 \cdot \log_2(0.2) + (-0.8) \cdot \log_2(0.8) \\ &= -0.2 \cdot -2.321 + (-0.8) \cdot -0.3219 \\ &= 0.4642 + 0.2575 \\ &= 0.7217 \end{aligned}$$

➤ Perhitungan gain outlook

$$\begin{aligned} \text{Gain: Entropy (total)} &= \sum_{i=1}^n \frac{|\text{outlook}_i|}{|\text{total}|} \times \text{Entropy (outlook}_i) \\ &= 0.8631 - \left(\left(\frac{5}{12} \times 0.9707 \right) + \left(\frac{4}{12} \times 0 \right) + \left(\frac{3}{12} \times 0.7217 \right) \right) \\ &= 0.8631 - (0.3966 + 0 + 0.2577) \\ &= 0.8631 - 0.6543 \\ &= 0.2088 \end{aligned}$$

➤ entropy temperatur

1. hot

$$\begin{aligned} \text{Entropy (hot)} &= \left(-\frac{2}{4} \times \log_2 \left(\frac{2}{4} \right) \right) + \left(-\frac{2}{4} \times \log_2 \left(\frac{2}{4} \right) \right) \\ &= -0.5 \cdot \log_2(0.5) + (-0.5) \cdot \log_2(0.5) \\ &= -0.5 \cdot -1 + (-0.5) \cdot -1 \\ &= 0.5 + 0.5 = 1 \end{aligned}$$

2. Mid

$$\begin{aligned}
 \text{Entropy (Mid)} &= \left(-\frac{2}{6} \times \log_2\left(\frac{2}{6}\right)\right) + \left(-\frac{4}{6} \times \log_2\left(\frac{4}{6}\right)\right) \\
 &= -0,333 \cdot \log_2(0,333) + (-0,666) \cdot \log_2(0,666) \\
 &= -0,333 - 1,5863 + (-0,666) - 0,5863 \\
 &= 0,5282 + 0,3964 \\
 &= 0,9186
 \end{aligned}$$

3. Cool

$$\begin{aligned}
 \text{entropy (Cool)} &= \left(-\frac{0}{4} \times \log_2\left(\frac{0}{4}\right)\right) + \left(-\frac{4}{4} \times \log_2\left(\frac{4}{4}\right)\right) \\
 &= 0 \cdot \log_2 0 + (-1) \cdot \log_2 1 \\
 &= 0 \cdot 0 + (-1) \cdot 0 = 0
 \end{aligned}$$

→ Perhitungan Gain temperature

$$\begin{aligned}
 \text{Gain} &= 0,8631 - \left(\left(\frac{4}{10} \times 1\right) + \left(\frac{6}{10} \times 0,9186\right) + \left(\frac{4}{10} \times 0\right)\right) \\
 &= 0,8631 - (0,2857 + 0,3936 + 0) \\
 &= 0,8631 - 0,6793 = 0,1838
 \end{aligned}$$

→ Perhitungan Entropy humidity

1. high

$$\begin{aligned}
 \text{Entropy (high)} &= \left(-\frac{4}{7} \times \log_2\left(\frac{4}{7}\right)\right) + \left(-\frac{3}{7} \times \log_2\left(\frac{3}{7}\right)\right) \\
 &= -0,5714 \cdot \log_2(0,5714) + (-0,4285) \cdot \log_2(0,4285) \\
 &= -0,5714 \cdot -0,8073 + (-0,4285) \cdot -1,225 \\
 &= 0,4612 + 0,5238 \\
 &= 0,985
 \end{aligned}$$

2. Normal

$$\begin{aligned}
 \text{entropy (normal)} &= \left(-\frac{0}{7} \log_2\left(\frac{0}{7}\right)\right) + \left(-\frac{7}{7} \times \log_2\left(\frac{7}{7}\right)\right) \\
 &= 0 \cdot \log_2 0 + (-1) \cdot \log_2 1 \\
 &= 0 \cdot 0 + (-1) \cdot 0 = 0
 \end{aligned}$$

→ Perhitungan Gain Humidity

$$\begin{aligned}
 \text{Gain} &= 0,8631 - \left(\left(\frac{7}{10} \times 0,985\right) + \left(\frac{3}{10} \times 0\right)\right) \\
 &= 0,8631 - (0,4925 + 0) \\
 &= 0,8631 - 0,4925 = 0,3706
 \end{aligned}$$

→ Entropy windy

1. No

$$\begin{aligned} \text{Entropy (No)} &= \left(-\frac{2}{8} \log_2 \left(\frac{2}{8}\right)\right) + \left(-\frac{6}{8} \log_2 \left(\frac{6}{8}\right)\right) \\ &= -0,25 \log_2 (0,25) + (-0,75) \log_2 (0,75) \\ &= -0,25 \cdot -2 + (-0,75) \cdot -0,4149 \\ &= 0,5 + 0,3111 = 0,8111 \end{aligned}$$

2. Yes

$$\begin{aligned} \text{Entropy (Yes)} &= \left(-\frac{2}{6} \log_2 \left(\frac{2}{6}\right)\right) + \left(-\frac{4}{6} \log_2 \left(\frac{4}{6}\right)\right) \\ &= -0,333 \log_2 (0,333) + (-0,666) \log_2 (0,666) \\ &= -0,333 \cdot (-1,5863) + (-0,666) \cdot (-0,863) \\ &= 0,5282 + 0,5794 = 1,1076 \end{aligned}$$

→ Gain windy

$$\begin{aligned} \text{Gain} &= 0,8631 - \left(\left(\frac{2}{8} \times 0,8111\right) + \left(\frac{6}{8} \times 1,1076\right)\right) \\ &= 0,8631 - (0,2528 + 0,9054) \\ &= 0,8631 - 1,1582 = -0,2951 \end{aligned}$$

→ Jadi hasil dikebutkan atribut dgn Gain tertinggi adalah Humidity yaitu 0,3706. Humidity memiliki 2 atribut yaitu High dan normal. nilai atribut normal sudah mengklasifikasi menjadi 1 keputusan play.

(Humidity)

High / normal
☐ ? ☒ yes

2 Node 2

Node 2		Smk Base	Don't play	Play	Entropy	Gain
Humidity High		7	4	3	0,985	
outlook						0,6993
	sunny	3	3	0	0	
	cloud	2	0	2	0	
	Rainy	2	1	1	1	
temperature						0,02
	Hot	3	2	1	0,9186	
	mid	4	2	2	1	
	cool	0	0	0	0	
windy	yes	4	2	2	1	
	no	3	2	1	0,9186	

$$\begin{aligned}
 \rightarrow \text{Entropy (total)} &= \left(-\frac{4}{7} \times \log_2 \left(\frac{4}{7} \right) \right) \\
 &\quad \left(-\frac{2}{7} \times \log_2 \left(\frac{2}{7} \right) \right) \\
 &= -0,5714 \cdot \log_2(0,5714) + (-0,428) \log_2(0,428) \\
 &= -0,5714 - 0,8073 + (-0,428) - 1,225 \\
 &= 0,4612 + 0,5238 = 0,985
 \end{aligned}$$

\rightarrow Entropy outlook high

$$\begin{aligned}
 1. \text{ Sunny} &= \left(-\frac{3}{3} \times \log_2 \left(\frac{3}{3} \right) \right) + \left(-\frac{0}{3} \times \log_2 \left(\frac{0}{3} \right) \right) \\
 &= (-1) \cdot \log_2 1 + 0 \cdot \log_2 0 \\
 &= (-1) \cdot 0 + 0 \cdot 0 = 0
 \end{aligned}$$

$$\begin{aligned}
 2. \text{ Cloudy} &= \left(-\frac{0}{2} \times \log_2 \left(\frac{0}{2} \right) \right) + \left(-\frac{2}{2} \times \log_2 \left(\frac{2}{2} \right) \right) \\
 &= 0 \cdot \log_2 0 + (-1) \log_2 1 \\
 &= 0 \cdot 0 + (-1) 0 = 0
 \end{aligned}$$

$$\begin{aligned}
 3. \text{ Rainy} &= \left(-\frac{1}{2} \times \log_2 \left(\frac{1}{2} \right) \right) + \left(-\frac{1}{2} \times \log_2 \left(\frac{1}{2} \right) \right) \\
 &= -0,5 \cdot \log_2(0,5) + (-0,5) \cdot \log_2(0,5) \\
 &= -0,5 \cdot -1 + (-0,5) \cdot -1 \\
 &= 0,5 + 0,5 = 1
 \end{aligned}$$

\rightarrow Gain outlook high

$$\begin{aligned}
 \text{Gain} &= 0,985 - \left(\left(\frac{3}{7} \times 0 \right) + \left(\frac{2}{7} \times 0 \right) + \left(\frac{2}{7} \times 1 \right) \right) \\
 &= 0,985 - (0 + 0 + 0,2857) \\
 &= 0,985 - 0,2857 = 0,6993
 \end{aligned}$$

\rightarrow Entropy temperature high

$$\begin{aligned}
 1. \text{ Hot} &= \left(-\frac{2}{3} \cdot \log_2 \left(\frac{2}{3} \right) \right) + \left(-\frac{1}{3} \times \log_2 \left(\frac{1}{3} \right) \right) \\
 &= -0,666 \cdot \log_2(0,666) + (-0,333) \cdot \log_2(0,333) \\
 &= -0,666 - 0,5863 + (-0,333) \cdot -1,5863 \\
 &= 0,9186
 \end{aligned}$$

$$\begin{aligned}
 2. \text{ mild} &= \left(-\frac{2}{4} \times \log_2 \left(\frac{2}{4} \right) \right) + \left(-\frac{2}{4} \times \log_2 \left(\frac{2}{4} \right) \right) \\
 &= -0,5 \cdot \log_2(0,5) + (-0,5) \cdot \log_2(0,5) \\
 &= -0,5 \cdot -1 + (-0,5) \cdot -1 = 0,5 + 0,5 = 1
 \end{aligned}$$

$$\begin{aligned}
 3. \text{ cool} &= \left(-\frac{0}{0} \times \log_2 \left(\frac{0}{0} \right) \right) + \left(-\frac{0}{0} \times \log_2 \left(\frac{0}{0} \right) \right) \\
 &= 0
 \end{aligned}$$

→ Gain temperature High

$$\begin{aligned} \text{Gain} &= 0,985 - \left(\left(\frac{3}{7} \times 0,9186 \right) + \left(\frac{4}{7} \times 1 \right) + \left(\frac{0}{7} \times 0 \right) \right) \\ &= 0,985 - (0,3936 + 0,5714 + 0) \\ &= 0,02 \end{aligned}$$

→ Entropy windy High

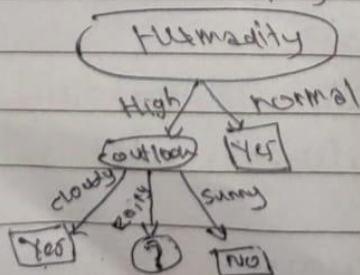
$$\begin{aligned} 1. \text{No} &= \left(-\frac{2}{4} \times \log_2 \left(\frac{2}{4} \right) \right) + \left(-\frac{2}{4} \times \log_2 \left(\frac{2}{4} \right) \right) \\ &= -0,5 \cdot \log_2 (0,5) + (-0,5) \cdot \log_2 (0,5) \\ &= -0,5 \cdot -1 + (-0,5) \cdot -1 \\ &= 1 \end{aligned}$$

$$\begin{aligned} 2. \text{Yes} &= \left(-\frac{2}{3} \times \log_2 \left(\frac{2}{3} \right) \right) + \left(-\frac{1}{3} \times \log_2 \left(\frac{1}{3} \right) \right) \\ &= -0,666 \cdot \log_2 (0,666) + (-0,333) \cdot \log_2 (0,333) \\ &= -0,666 - 0,5863 + (-0,333) - 1,5863 \\ &= 0,9186 \end{aligned}$$

→ Gain windy High

$$\begin{aligned} \text{Gain} &= 0,985 - \left(\left(\frac{4}{7} \times 1 \right) + \left(\frac{3}{7} \times 0,9186 \right) \right) \\ &= 0,985 - (0,5714 + 0,3936) \\ &= 0,985 - 0,965 = 0,02 \end{aligned}$$

Jadi, gain tertinggi adalah Outlook yaitu 0,6993. Maka outlook menjadi cabang attribute high. Dari attribute outlook memiliki 3 nilai yaitu Sunny dgn klasifikasi Karur 1 (Don't play), cloudy dgn klasifikasi Karur 1 (play) dan Rainy masih perlu dihitung



→ Node 3

Node 3

	Don't play	Play	entropy	Gain
Humidity High and outlook rainy Temperature	2	1	1	
Hot	0	0	0	0
mild	2	1	1	
cool	0	0	0	
windy				1
No	1	1	0	
Yes	1	0	0	

→ entropy rainy

$$1. \text{ Hot} = \left(\frac{0}{0} \times \log_2 \left(\frac{0}{0} \right) \right) + \left(\frac{0}{0} \times \log_2 \left(\frac{0}{0} \right) \right)$$

$$= 0$$

$$2. \text{ Mild} = \left(\frac{1}{2} \times \log_2 \left(\frac{1}{2} \right) \right) + \left(\frac{1}{2} \times \log_2 \left(\frac{1}{2} \right) \right)$$

$$= -0,5 \cdot \log_2 (0,5) + (-0,5) \cdot \log_2 (0,5)$$

$$= -0,5 \cdot (-1) + (-0,5) \cdot (-1)$$

$$= 1$$

$$3. \text{ cool} = \left(\frac{0}{0} \times \log_2 \left(\frac{0}{0} \right) \right) + \left(\frac{0}{0} \times \log_2 \left(\frac{0}{0} \right) \right)$$

$$= 0$$

→ Perhitungan entropy windy rainy

$$1. \text{ No} = \left(\frac{0}{1} \times \log_2 \left(\frac{0}{1} \right) \right) + \left(\frac{1}{1} \times \log_2 \left(\frac{1}{1} \right) \right)$$

$$= 0 \cdot \log_2 0 + (-1) \log_2 (1)$$

$$= 0$$

$$2. \text{ Yes} = \left(\frac{1}{1} \times \log_2 \left(\frac{1}{1} \right) \right) + \left(\frac{0}{1} \times \log_2 \left(\frac{0}{1} \right) \right)$$

$$= -1 \log_2 (1) + 0 \cdot \log_2 0$$

$$= 0$$

→ Gain windy Rainy

$$\text{Gain} = 1 - \left(\left(\frac{1}{2} \times 0 \right) + \left(\frac{1}{2} \times 0 \right) \right)$$

$$= 1 - (0 + 0)$$

$$= 0$$

Jadi Gain tertinggi adalah windy yaitu 1. sehingga windy
menjadi cabang dari atribut rainy. Ada 2 nilai yaitu noden yes
nilai atribut no sudah mengklasifikasi kelas menjadi 1 (Play) dan nilai
atribut yes mengklasifikasi menjadi 1 (Don't Play). Sehingga akan
dihitung lagi.

