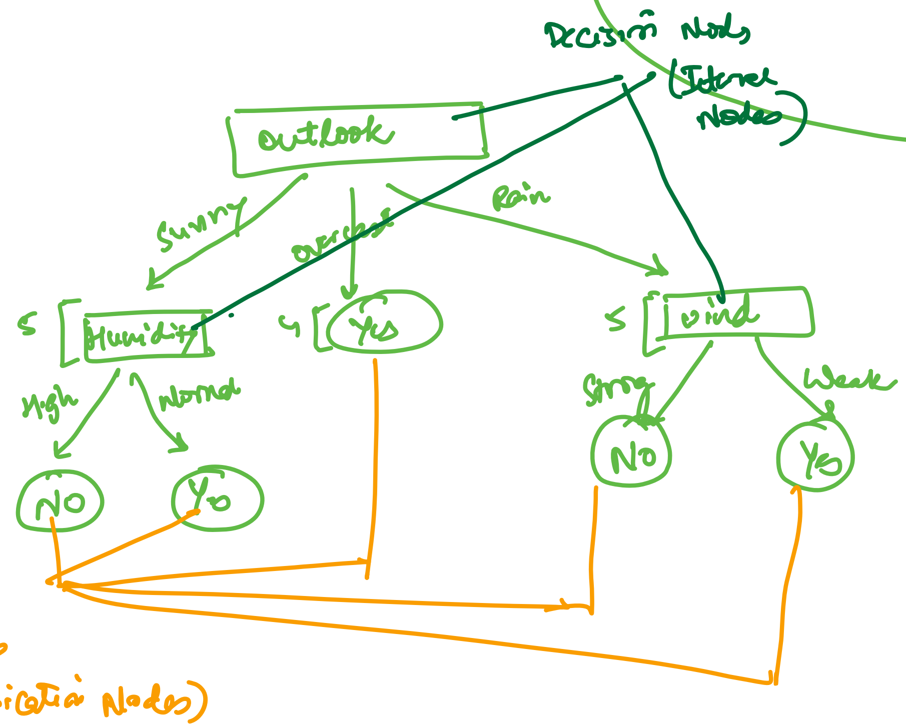
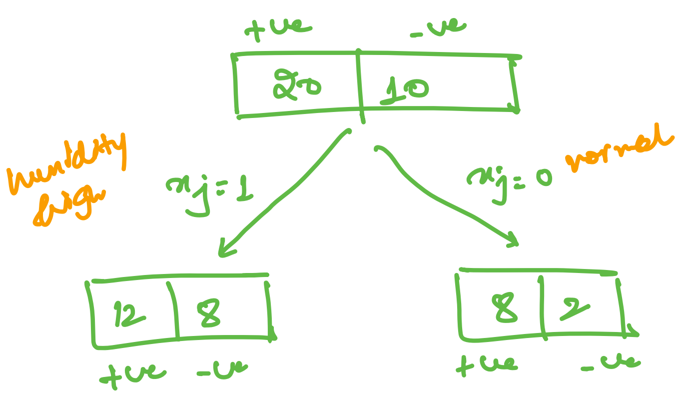
Decision Trees

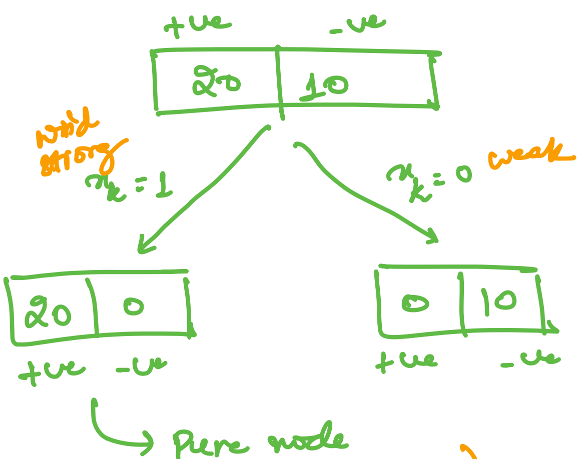
↳ Supervised Algo
↳ Classification & Regression

LR
LR
KNN
NaiveBayes

Day	Outlook	Temp	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







high entropy

min entropy ✓

Entropy: Degree of randomness

$$y \in \{1 \ldots \ldots x\}$$

$$P(y=k) = p_k$$

$$H(y) = -\sum_{k=1}^{r} p_k \log p_k$$

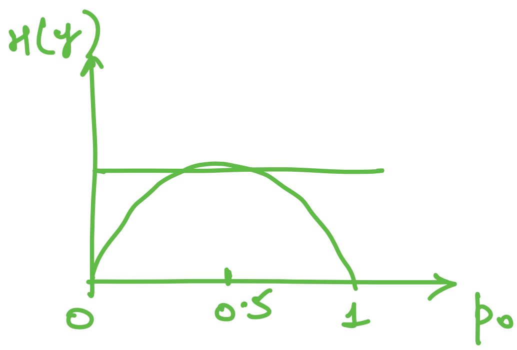
Boolean case: $y \in \{0, 1\}$ $\qquad p_0 + p_1 = 1$

$$H(y) = -\left(p_0 \log p_0 + p_1 \log p_1\right)$$

$$H(y) = -\left(\underbrace{p_0 \log p_0}_{0} + \underbrace{(1-p_0)}_{1}\underbrace{\log p_1}_{0}\right)$$

$p_0 = 0$
$p_1 = 1$

$= 0$



$$H(y) = -\left(p_0 \log p_0 + (1-p_0)\log(1-p_0)\right)$$

$$\frac{dH}{dp_0} = -\left(\log p_0 + \frac{p_0}{p_0} - \log(1-p_0) - \frac{(1-p_0)}{(1-p_0)}\right)$$

$$= -\left(\log p_0 + \cancel{1} - \log(1-p_0) - \cancel{1}\right) = 0$$

$$\log p_0 = \log(1-p_0)$$

$$p_0 = 1 - p_0$$

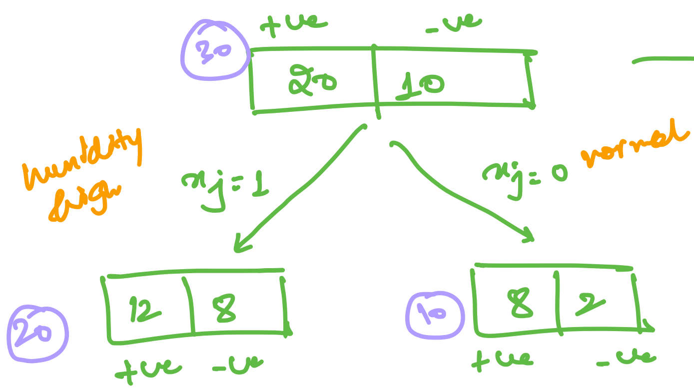
$$2p_0 = 1$$

$$p_0 = 0.5 = 1/2$$

$y \in \{1, 2, 3 \cdots r\}$

$H(y) = -\sum_{k=1}^{r} p_k \log p_k$ is max at $p_k = 1/r \ \forall k$

Entropy before splitting



$$H(y) = -\left(\frac{20}{30}\log\frac{20}{30} + \frac{10}{30}\log\frac{10}{30}\right)$$

$$H(y|x_j=0) = -\left(\frac{8}{10}\log\frac{8}{10} + \frac{2}{10}\log\frac{2}{10}\right)$$

high entropy

$$H(y|x_j=1) = -\left(\frac{12}{20}\log\frac{12}{20} + \frac{8}{20}\log\frac{8}{20}\right)$$

Combine

$$\left.\begin{aligned} H(y|x_j) &= P(x_j=1)*H(y|x_j=1) + P(x_j=0)*H(y|x_j=0) \\ &= \frac{20}{30} \times -\left(\frac{12}{20}\log\frac{12}{20} + \frac{8}{20}\log\frac{8}{20}\right) + \\ &\quad \frac{10}{30} \times -\left(\frac{8}{10}\log\frac{8}{10} + \frac{2}{10}\log\frac{2}{10}\right) \end{aligned}\right\} \text{Entropy after splitting}$$

Entropy

1. pick after splitting min entropy   0.2 ✓   0.7
2. Mutual Information $= H(y) - H(y|x_j)$: Entropy before splitting - Entropy after splitting

   Select the feature which gives you max information gain

   0.8-0.2 = 0.6 ✓   0.8-0.7 = 0.1

Gini Index

$$Gini(y) = 1 - \sum_{k=1}^{r} p_k^2 \quad \Big\}\ \text{min}$$

Decision Tree Example:

Entropy (ID3 Algorithm)

← x → ← y →

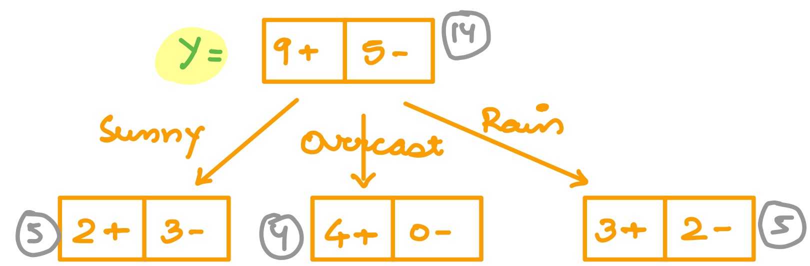
Day	Outlook	Temp	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes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

Find the attribute which is giving maximum information out of available attributes.

Root Node

Day	Outlook	Temp	Humidity	Wind	Play Tennis
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



$$H(Y) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.94$$ Initial Entropy (Before split)

$$H(Y|x_j = \text{Sunny}) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = 0.971$$

$$H(Y|x_j = \text{Overcast}) = -\frac{4}{4}\log_2\frac{4}{4} - \frac{0}{4}\log_2\frac{0}{4} = 0$$

$$H(Y|x_j = \text{Rain}) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.971$$

$$MI(Y, \overset{\text{outlook}}{x_j}) = \underbrace{H(Y)}_{\text{Before split}} - H(Y|x_j)$$

$$MI(y, x_j = \text{outlook}) = H(y) - \Big(P(x_j = \text{Sunny}) \cdot H(y \mid x_j = \text{Sunny}) + P(x_j = \text{Overcast}) \cdot H(y \mid x_j = \text{overcast}) + P(x_j = \text{Rain}) \cdot H(y \mid x_j = \text{Rain})\Big)$$

After split

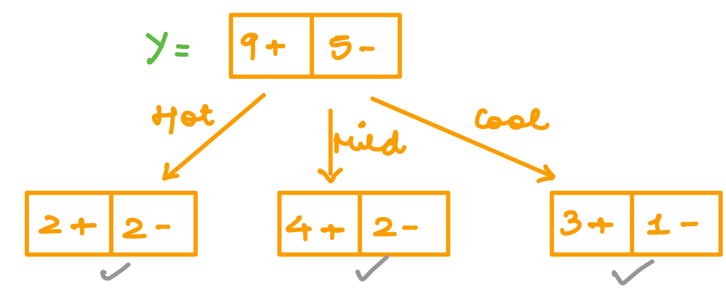
$$= 0.94 - \frac{5}{14}(0.971) - \frac{4}{14} \cdot 0 - \frac{5}{14}(0.971)$$

$$= 0.2464$$

Information Gain in outlook Attribute
(Mutual Information)

Attribute : Temp

Day	Outlook	Temp	Humidity	Wind	Play Tennis
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



$$H(y) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.94$$

$$H(y \mid x_j = \text{Hot}) = -\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4} = 1$$

$$H(y \mid x_j = \text{Mild}) = -\frac{4}{6}\log\frac{4}{6} - \frac{2}{6}\log_2\frac{2}{6} = 0.9183$$

$$H(y \mid x_j = \text{Cool}) = -\frac{3}{4}\log_2\frac{3}{4} - \frac{1}{4}\log_2\frac{1}{4} = 0.8113$$

$$MI(y \mid x_j = \text{temp}) = H(y) - \boxed{\Big(P(x_j = \text{Hot}) \cdot H(y \mid x_j = \text{Hot}) + P(x_j = \text{Mild}) \cdot H(y \mid x_j = \text{Mild}) + P(x_j = \text{Cool}) \cdot H(y \mid x_j = \text{Cool})\Big)}$$

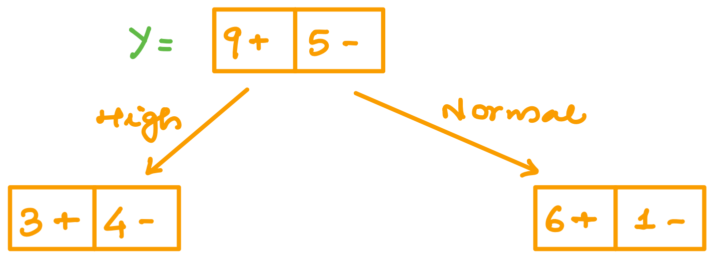
weighted avg of entropy after split

$$= 0.94 - \frac{4}{14}(1) - \frac{6}{14}(0.9183) - \frac{4}{14}(0.8113)$$

$$= 0.0289$$

Attribute: Humidity

Day	Outlook	Temp	Humidity	Wind	Play Tennis
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



$$H(Y) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.94$$

$$H(Y|x_j = high) = -\frac{3}{7}\log_2\frac{3}{7} - \frac{4}{7}\log_2\frac{4}{7}$$

$$= 0.9852$$

$$H(Y|x_j = normal) = -\frac{6}{7}\log_2\frac{6}{7} - \frac{1}{7}\log_2\frac{1}{7}$$

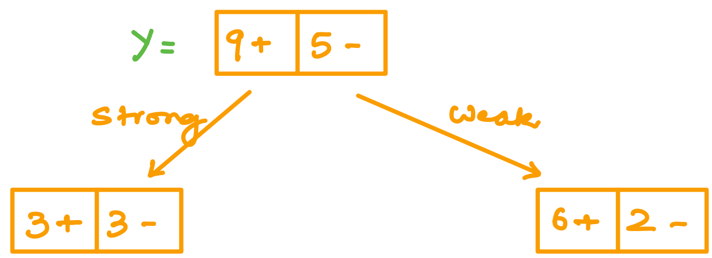
$$= 0.5916$$

$$MI(Y|x_j = humidity) = 0.94 - \frac{7}{14}(0.9852) - \frac{7}{14}(0.5916)$$

$$= 0.1516$$

Attribute: Wind

Day	Outlook	Temp	Humidity	Wind	Play Tennis
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



$$H(Y) = -\frac{9}{14}\log_2\frac{9}{14} - \frac{5}{14}\log_2\frac{5}{14} = 0.94$$

$$H(Y|x_j = strong) = 1.0$$

$$H(Y|x_j = weak) = -\frac{6}{8}\log_2\frac{6}{8} - \frac{2}{8}\log_2\frac{2}{8}$$

$$= 0.8113$$

$$MI(Y|x_j = wind) = 0.94 - \frac{6}{14}(1.0) - \frac{8}{14}(0.8113) = 0.0478$$

$MI(y \mid x_j : \text{outlook}) = 0.2464$

$MI(y \mid x_j = \text{temp}) = 0.0289$

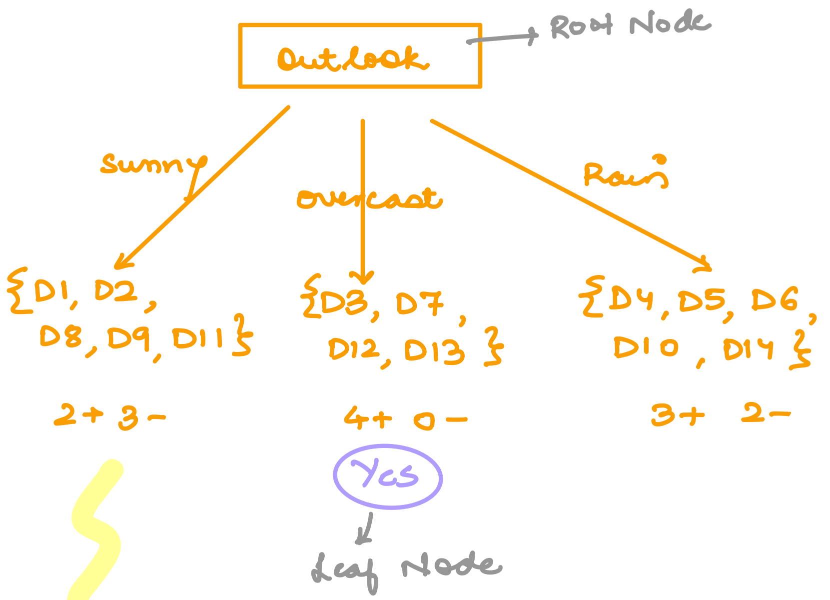
$MI(y \mid x_j = \text{humidity}) : 0.1516$

$MI(y \mid x_j = \text{wind}) = 0.0478$

which attribute has maximum MI (information gain)?

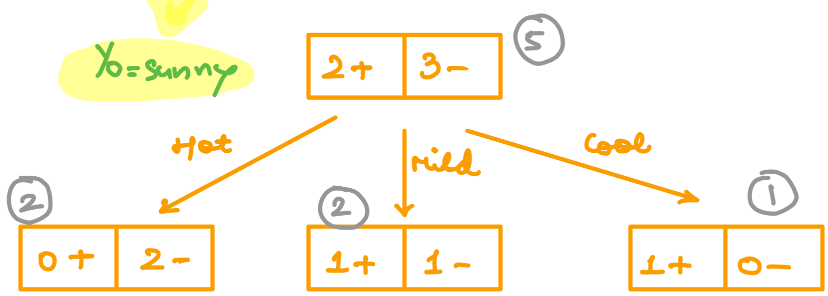
Take outlook as root node

Day	Outlook	Temp	Humidity	Wind	Play Tennis
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: Temp

Day	Temp	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



$H(Y_o : \text{sunny}) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = 0.97$

$H(Y_o : \text{sunny} \mid x_j : \text{hot}) : 0$

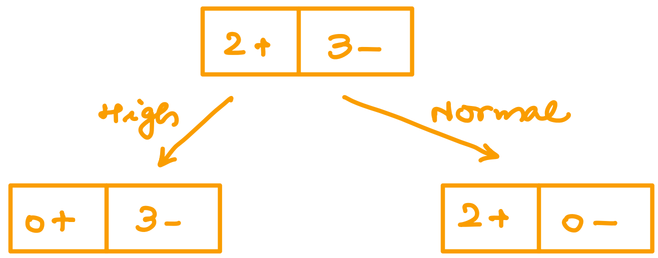
$H(Y_o : \text{sunny} \mid x_j = \text{mild}) : 1$

$H(Y_o : \text{sunny} \mid x_j : \text{cool}) = 0$

$$MI(Y_{o=outlook} | x_j = temp) = 0.97 - \frac{2}{5}(0) - \frac{2}{5}(1) - \frac{1}{5}(0) = 0.570$$

Attribute: Humidity

Day	Temp	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



$$H(Y_{o=sunny}) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = 0.97$$

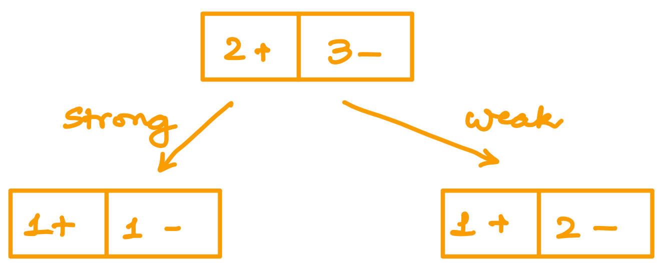
$$H(Y_{o=sunny} | x_j = high) = 0.0$$

$$H(Y_{o=sunny} | x_j = normal) = 0.0$$

$$MI(Y_{o=sunny} | x_j = humidity) = 0.97 - \frac{3}{5}(0) - \frac{2}{5}(0) = 0.97$$

Attribute: wind

Day	Temp	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



$$H(Y) = -\frac{2}{5}\log_2\frac{2}{5} - \frac{3}{5}\log_2\frac{3}{5} = 0.97$$

$$H(Y | x_j = strong) = 1.0$$

$$H(Y | x_j = weak) = -\frac{1}{3}\log_2\frac{1}{3} - \frac{2}{3}\log_2\frac{2}{3} = 0.9183$$

$$MI(Y | x_j = wind) = 0.97 - \frac{2}{5}(1) - \frac{3}{5}(0.9183) = 0.0192$$

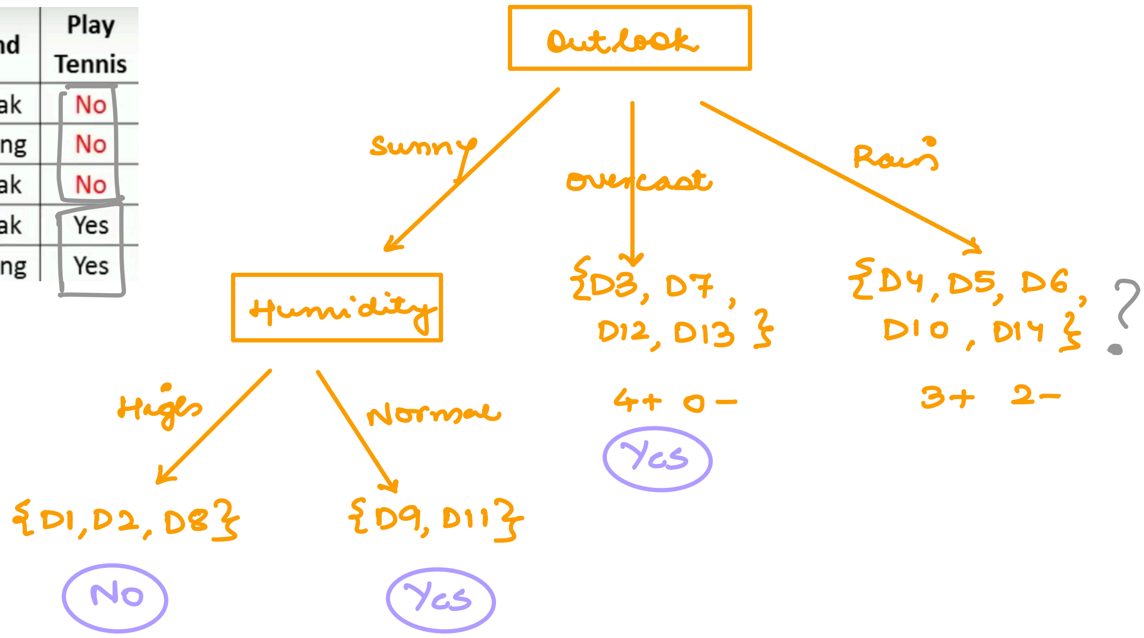
$MI(Y|x_j: temp) = 0.570$

$MI(Y|x_j: humidity) = 0.97$

$MI(Y|x_j: wind) = 0.0192$

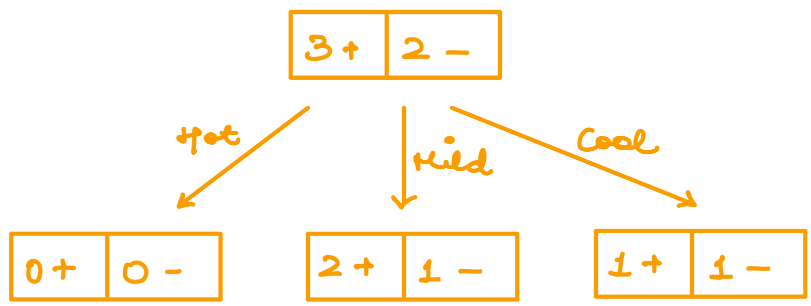
Maximum information gain is through humidity.

Day	Temp	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: temp

Day	Temp	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



$$H(Y) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.97$$

$$H(Y|x_j = hot) = 0$$

$$H(Y|x_j = mild) = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} = 0.9183$$

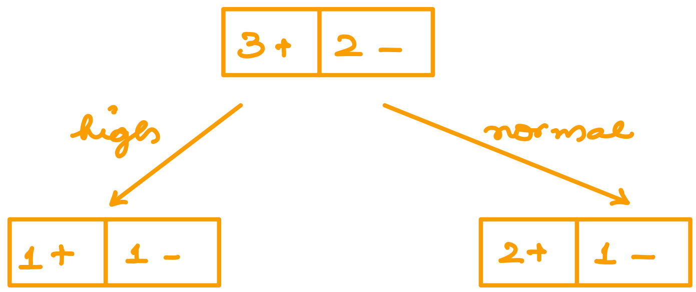
$$H(Y|x_j: cool) = 1.0$$

$$MI(y|x_j = temp) = 0.97 - \frac{0}{5}(0) - \frac{3}{5}(0.918) - \frac{2}{5}(1.0)$$

$$= 0.0192$$

Attribute: humidity

Day	Temp	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



$$H(y) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.97$$

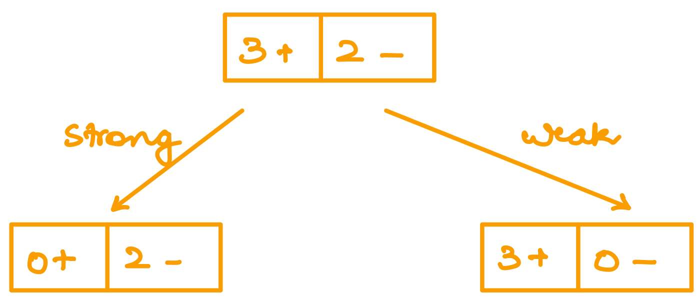
$$H(y|x_j = high) = 1.0$$

$$H(y|x_j = normal) = -\frac{2}{3}\log_2\frac{2}{3} - \frac{1}{3}\log_2\frac{1}{3} = 0.9183$$

$$MI(y|x_j = humidity) = 0.97 - \frac{2}{5}(1.0) - \frac{3}{5}(0.9183) = 0.0192$$

Attribute: Wind

Day	Temp	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



$$H(y) = -\frac{3}{5}\log_2\frac{3}{5} - \frac{2}{5}\log_2\frac{2}{5} = 0.97$$

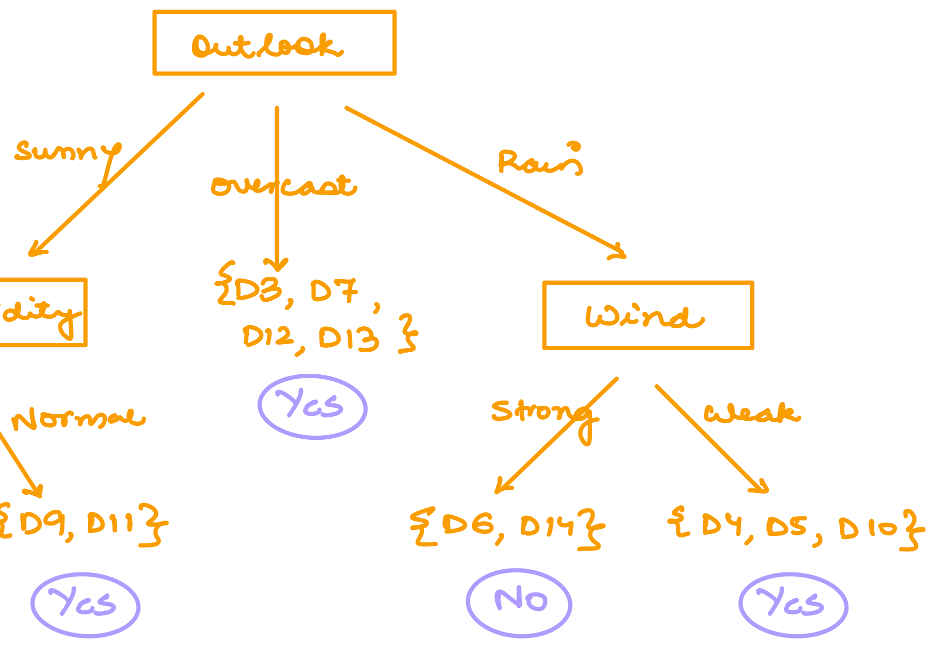
$$H(y|x_j = strong) = 0$$

$$H(y|x_j = weak) = 0$$

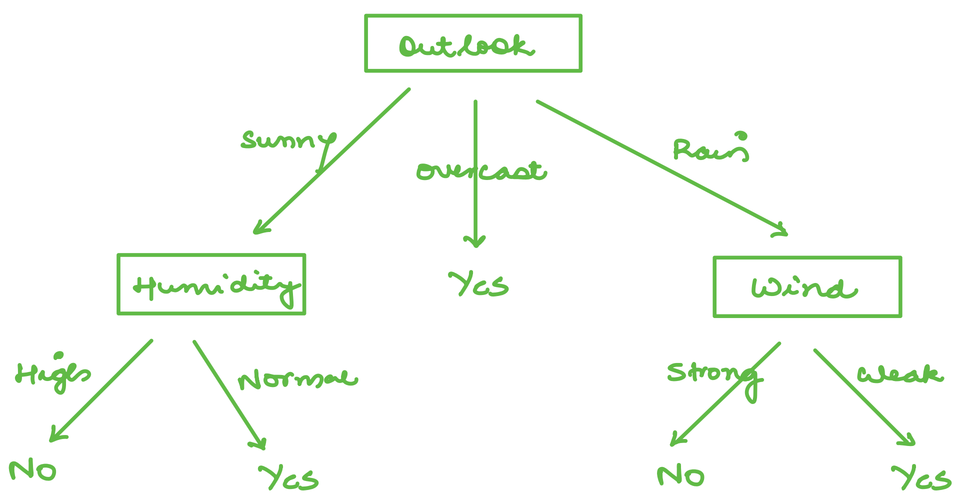
$$MI(y|x_j: wind) = 0.97 - \frac{2}{5}(0) - \frac{3}{5}(0) = 0.97$$

$$\left.\begin{aligned} MI(y|x_j = temp) &= 0.0192 \\ MI(y|x_j : humidity) &= 0.0192 \\ MI(y|x_j : wind) &= 0.97 \end{aligned}\right\} \text{maximum for wind}$$

Day	Temp	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



Final Decision Tree:



Decision Tree using Gini Index

← x → ← y →

Weekend	Weather	Parents	Money	Decision
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Stay In
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis

Calculate Gini Index of every attribute, attribute having minimum gini index will be selected.

$$Gini(y) = 1 - \left[\left(\overset{\text{Cinema}}{\frac{6}{10}}\right)^2 + \left(\overset{\text{Tennis}}{\frac{2}{10}}\right)^2 + \left(\overset{\text{Stay In}}{\frac{1}{10}}\right)^2 + \left(\underset{\text{Shopping}}{\frac{1}{10}}\right)^2 \right]$$

$$= 0.58$$

Attribute : Weather

Weekend	Weather	Parents	Money	Decision
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Stay In
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis

$$Gini(y \mid x_j = Sunny) = 1 - \left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 \right] = 0.444$$

$$Gini(y \mid x_j = Rainy) = 1 - \left[\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 \right] = 0.444$$

$$Gini(y \mid x_j = windy) = 1 - \left[\left(\frac{3}{4}\right)^2 + \left(\frac{1}{4}\right)^2 \right] = 0.375$$

$$Gini(y \mid x_j = weather) = \frac{3}{10} * 0.444 + \frac{3}{10} * 0.444 + \frac{4}{10} * 0.375$$

$$= 0.416$$

Attribute: Parents

Weekend	Weather	Parents	Money	Decision
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Stay In
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis

6 Cinema
2 Tennis
1 Stay In
1 Shopping
10

Yes:
5 Cinema
0 Tennis
0 Stay In
0 Shopping
5

No:
1 Cinema
2 Tennis
1 Stay In
1 Shopping
5

$$Gini(y \mid x_j = Yes) = 1 - \left[\left(\frac{5}{5}\right)^2\right] = 0$$

$$Gini(y \mid x_j = No) = 1 - \left[\left(\frac{1}{5}\right)^2 + \left(\frac{2}{5}\right)^2 + \left(\frac{1}{5}\right)^2 + \left(\frac{1}{5}\right)^2\right] = 0.72$$

$$Gini(y \mid x_j = parents) = \frac{5}{10} * 0 + \frac{5}{10} * 0.72 = 0.36$$

Attribute: Money

Weekend	Weather	Parents	Money	Decision
W1	Sunny	Yes	Rich	Cinema
W2	Sunny	No	Rich	Tennis
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W5	Rainy	No	Rich	Stay In
W6	Rainy	Yes	Poor	Cinema
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W9	Windy	Yes	Rich	Cinema
W10	Sunny	No	Rich	Tennis

6 Cinema
2 Tennis
1 Stay In
1 Shopping

Poor:
3 Cinema
0 Tennis
0 Stay In
0 Shopping

Rich:
3 Cinema
2 Tennis
1 Stay In
1 Shopping

$$Gini(y \mid x_j = Poor) = 1 - \left[\left(\frac{3}{3}\right)^2\right] = 0$$

$$Gini(y \mid x_j = Rich) = 1 - \left[\left(\frac{3}{7}\right)^2 + \left(\frac{2}{7}\right)^2 + \left(\frac{1}{7}\right)^2 + \left(\frac{1}{7}\right)^2\right] = 0.694$$

$$Gini(y \mid x_j = money) = \frac{3}{10} * 0 + \frac{7}{10} * 0.694 = 0.486$$

$$\left.\begin{aligned} Gini(y \mid x_j = weather) &= 0.416 \\ Gini(y \mid x_j = parents) &= 0.36 \\ Gini(y \mid x_j = money) &= 0.486 \end{aligned}\right\}$$

Gini for parents is minimum. We select parent as root node.

$\boxed{\text{Parents}}$

Yes

Weekend	Weather	Parents	Money	Decision
W1	Sunny	Yes	Rich	Cinema
W3	Windy	Yes	Rich	Cinema
W4	Rainy	Yes	Poor	Cinema
W6	Rainy	Yes	Poor	Cinema
W9	Windy	Yes	Rich	Cinema

Decision is always Cinema

No

Weekend	Weather	Parents	Money	Decision
W2	Sunny	No	Rich	Tennis
W5	Rainy	No	Rich	Stay In
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W10	Sunny	No	Rich	Tennis

?

Not pure ?

Attribute : weather

Weekend	Weather	Parents	Money	Decision
W2	Sunny	No	Rich	Tennis
W5	Rainy	No	Rich	Stay In
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W10	Sunny	No	Rich	Tennis

1 Cinema
2 Tennis
1 Stay In
1 Shopping

Sunny: 0 Cinema, 2 Tennis, 0 Stay In, 0 Shopping

Rainy: 0 Cinema, 0 Tennis, 1 Stay In, 0 Shopping

windy: 1 Cinema, 0 Tennis, 0 Stay In, 1 Shopping

$$Gini(y_{parents = No} \mid x_j = Sunny) = 1 - \left[\left(\frac{2}{2}\right)^2\right] = 0$$

$$Gini\left(y_{parents=NO} \mid x_j = Rainy\right) = 1 - \left[\left(\frac{1}{1}\right)^2\right] = 0$$

$$Gini\left(y_{parents=NO} \mid x_j = windy\right) = 1 - \left[\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2\right] = 0.5$$

$$Gini\left(y_{parents=NO} \mid x_j = weather\right) = \frac{2}{5} * 0 + \frac{1}{5} * 0 + \frac{2}{5} * 0.5 = 0.2$$

Attribute: Money

Weekend	Weather	Parents	Money	Decision
W2	Sunny	No	Rich	Tennis
W5	Rainy	No	Rich	Stay In
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W10	Sunny	No	Rich	Tennis

6 Cinema
2 Tennis
1 Stay In
1 Shopping

Poor:
1 Cinema
0 Tennis
0 StayIn
0 Shopping

Rich:
0 Cinema
2 Tennis
1 StayIn
1 Shopping

$$Gini\left(y_{parents=NO} \mid x_j = poor\right) = 1 - \left[\left(\frac{1}{1}\right)^2\right] = 0$$

$$Gini\left(y_{parents=NO} \mid x_j = rich\right) = 1 - \left[\left(\frac{2}{4}\right)^2 + \left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right)^2\right] = 0.625$$

$$Gini\left(y_{parents=NO} \mid x_j = money\right) = \frac{1}{5} * 0 + \frac{4}{5} * 0.625 = 0.5$$

$$\left.\begin{aligned} Gini\left(y_{parents=NO} \mid x_j = weather\right) &= 0.2 \\ Gini\left(y_{parents=NO} \mid x_j = money\right) &= 0.5 \end{aligned}\right\} \text{Weather is smaller.}$$

Parents

Yes

Cinema

No

Weekend	Weather	Parents	Money	Decision
W2	Sunny	No	Rich	Tennis
W5	Rainy	No	Rich	Stay In
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping
W10	Sunny	No	Rich	Tennis

Sunny

Weekend	Weather	Parents	Money	Decision
W2	Sunny	No	Rich	Tennis
W10	Sunny	No	Rich	Tennis

Rainy

Weekend	Weather	Parents	Money	Decision
W5	Rainy	No	Rich	Stay In

windy

Weekend	Weather	Parents	Money	Decision
W7	Windy	No	Poor	Cinema
W8	Windy	No	Rich	Shopping

?

Final Decision Tree:

Parents

Yes

No

Cinema

Weather

Sunny

Rainy

windy

Tennis

Stay In

Money

Rich

Poor

Shopping

Cinema

Decision Trees
(Discrete Attributes)

Entropy
(IG)
(MI)

Gini Index