Decision Trees Example Problem

Consider the following data, where the Y label is whether or not the child goes out to play.

Day	Weather	Temperature	Humidity	Wind	Play?
1	Sunny	Hot	High	Weak	No
2	Cloudy	Hot	High	Weak	Yes
3	Sunny	Mild	Normal	Strong	Yes
4	Cloudy	Mild	High	Strong	Yes
5	Rainy	Mild	High	Strong	No
6	Rainy	Cool	Normal	Strong	No
7	Rainy	Mild	High	Weak	Yes
8	Sunny	Hot	High	Strong	No
9	Cloudy	Hot	Normal	Weak	Yes
10	Rainy	Mild	High	Strong	No

Step 1: Calculate the IG (information gain) for each attribute (feature)

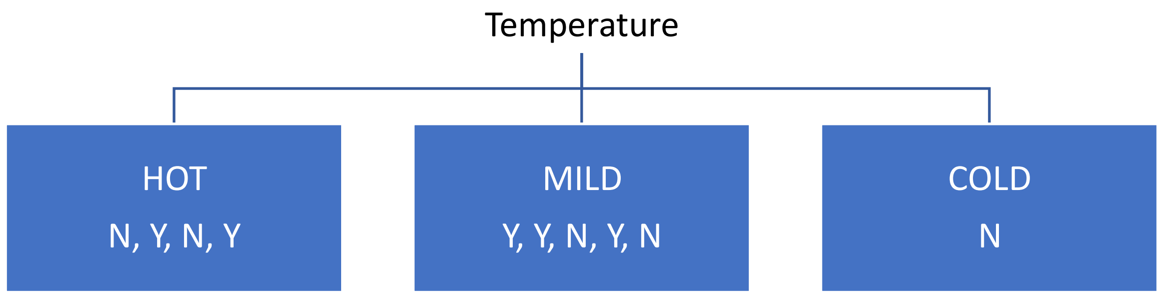
Initial entropy = $H(Y) = -\sum_y P(Y=y)\log_2 P(Y=y)$

$= -P(Y=yes)\log_2 P(Y=yes) - P(Y=no)\log_2 P(Y=no)$

$= -(0.5)\log_2(0.5) - (0.5)\log_2(0.5)$

$= 1$

Temperature:



Total entropy of this division is:

$$H(Y \mid temp) = -\sum_x P(temp = x) \sum_y P(Y=y \mid temp = x)\log_2 P(Y=y \mid temp = x)$$

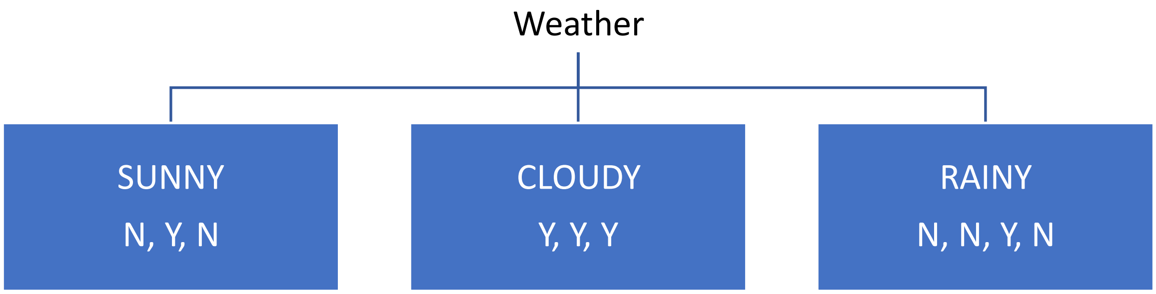
$$\begin{aligned} &= -\big(P(temp = H)\textstyle\sum_y P(Y=y \mid temp = H)\log_2 P(Y=y \mid temp = H) + \\ &\quad P(temp = M)\textstyle\sum_y P(Y=y \mid temp = M)\log_2 P(Y=y \mid temp = M) + \\ &\quad P(temp = C)\textstyle\sum_y P(Y=y \mid temp = C)\log_2 P(Y=y \mid temp = C)\big) \end{aligned}$$

$$\begin{aligned} &= -\Big((0.4)\big((\tfrac{1}{2})\log_2\left(\tfrac{1}{2}\right) + (\tfrac{1}{2})\log_2(\tfrac{1}{2})\big) + (0.5)\big(\left(\tfrac{3}{5}\right)\log_2\left(\tfrac{3}{5}\right) + \left(\tfrac{2}{5}\right)\log_2(\tfrac{2}{5})\big) + \\ &\quad (0.1)\big((1)\log_2(1) + (0)\log_2(0)\big)\Big) \end{aligned}$$

= 0.7884

IG(Y, temp) = 1 – 0.7884 = 0.2116

Weather:



Total entropy of this division is:

$$H(Y \mid weather) = -\sum_x P(weather = x) \sum_y P(Y = y \mid weather = x) \log_2 P(Y = y \mid weather = x)$$

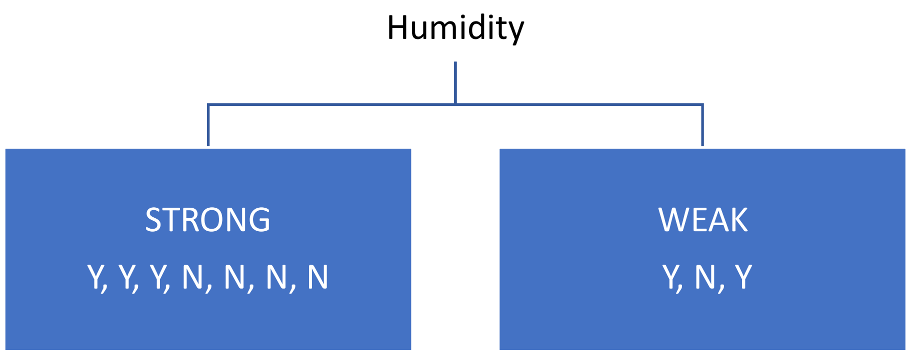
$$= -(P(weather = S) \textstyle\sum_y P(Y = y \mid weather = S) \log_2 P(Y = y \mid weather = S) + P(weather = C) \textstyle\sum_y P(Y = y \mid weather = C) \log_2 P(Y = y \mid weather = C) + P(weather = R) \textstyle\sum_y P(Y = y \mid weather = R) \log_2 P(Y = y \mid weather = R))$$

$$= -((0.3)((\tfrac{1}{3}) \log_2 \left(\tfrac{1}{3}\right) + (\tfrac{2}{3}) \log_2 (\tfrac{2}{3})) + (0.3)((1) \log_2 (1) + (0) \log_2 (0)) + (0.4)(\left(\tfrac{1}{4}\right) \log_2 \left(\tfrac{1}{4}\right) + \left(\tfrac{3}{4}\right) \log_2 (\tfrac{3}{4})))$$

= 0.6

IG(Y, weather) = 1 – 0.6 = 0.4

Humidity:



Total entropy of this division is:

$$H(Y \mid hum) = -\sum_{x} P(hum \ = x) \sum_{y} P(Y = y \mid hum = x) \log_2 P(Y = y \mid hum = x)$$

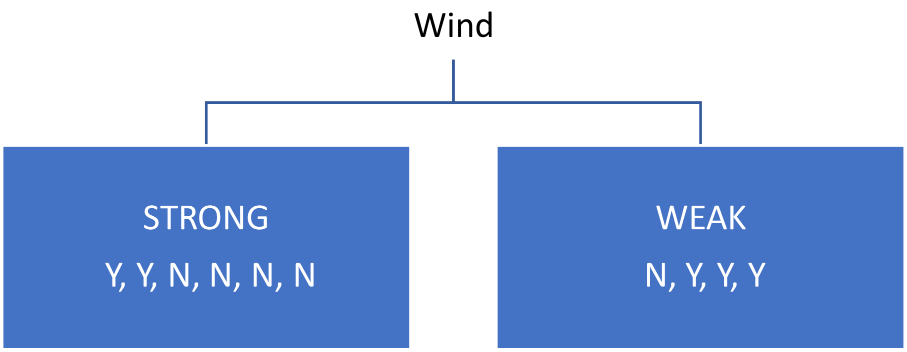
$$= -(P(hum \ = H) \textstyle\sum_y P(Y = y \mid hum = H) \log_2 P(Y = y \mid hum = H) \ + \\ P(hum = N) \textstyle\sum_y P(Y = y \mid hum = N) \log_2 P(Y = y \mid hum = N)$$

$$= -((0.7)((\tfrac{3}{7}) \log_2 \left(\tfrac{3}{7}\right) + (\tfrac{4}{7}) \log_2 (\tfrac{4}{7})) \ + (0.3)(\left(\tfrac{2}{3}\right) \log_2 \left(\tfrac{2}{3}\right) + \left(\tfrac{1}{3}\right) \log_2 (\tfrac{1}{3}))$$

= 0.8651

IG(Y, hum) = 1 – 0.8651 = 0.1349

Wind:



Total entropy of this division is:

$$H(Y \mid wind) = -\sum_{x} P(wind \ = x) \sum_{y} P(Y = y \mid wind = x) \log_2 P(Y = y \mid wind = x)$$

$$= -(P(wind \ = S) \textstyle\sum_y P(Y = y \mid wind = S) \log_2 P(Y = y \mid wind = S) \ +$$
$$P(wind = W) \textstyle\sum_y P(Y = y \mid wind = W) \log_2 P(Y = y \mid wind = W)$$

$$= -((0.6)((\tfrac{2}{6}) \log_2 \left(\tfrac{2}{6}\right) + (\tfrac{4}{6}) \log_2(\tfrac{4}{6})) \ + (0.4)(\left(\tfrac{1}{4}\right) \log_2 \left(\tfrac{1}{4}\right) + \left(\tfrac{3}{4}\right) \log_2(\tfrac{3}{4}))$$

= 0.8755

IG(Y, wind) = 1 – 0.8755 = 0.1245

Step 2: Choose which feature to split with!

IG(Y, wind) = 0.1245

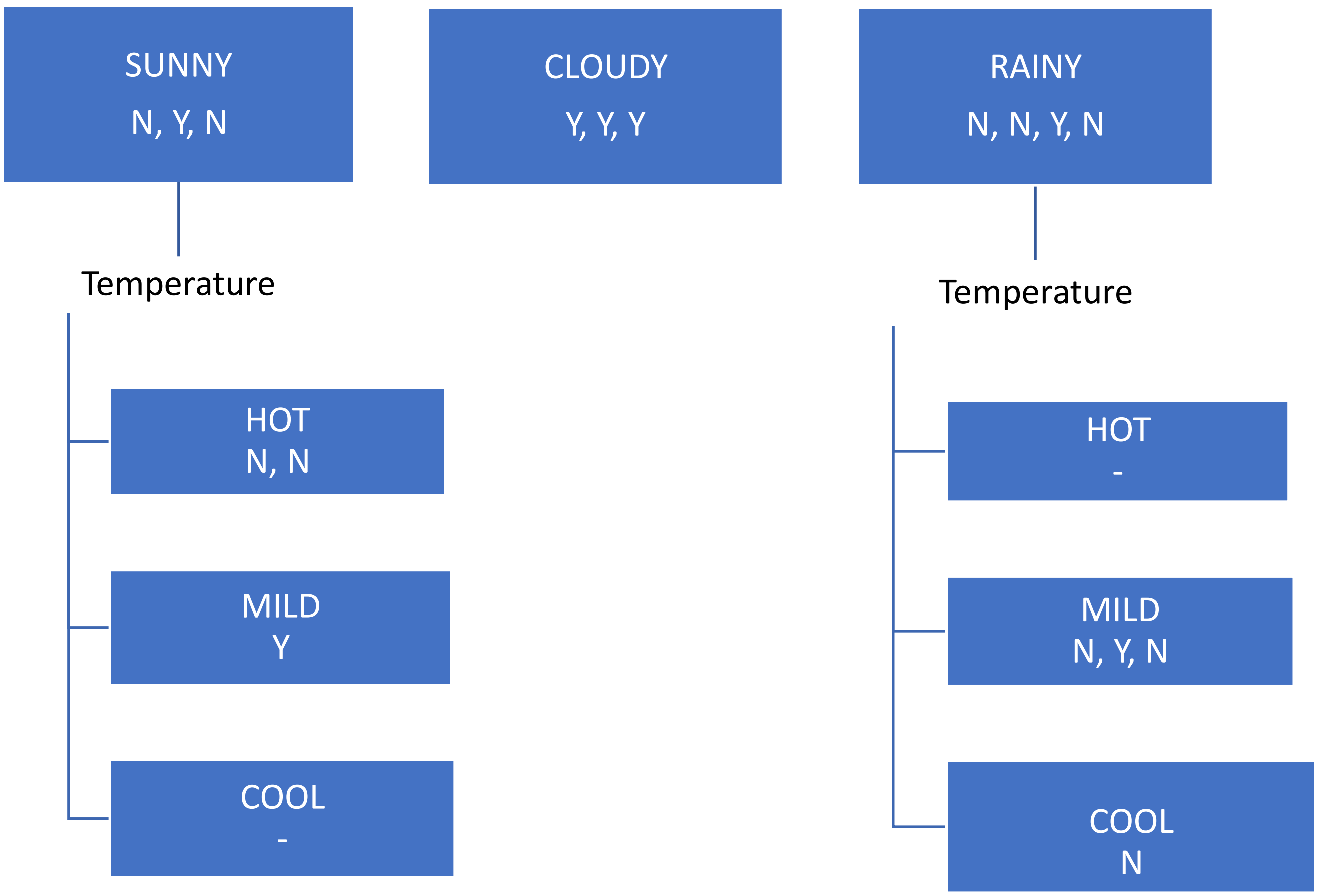
IG(Y, hum) = 0.1349

IG(Y, weather) = 0.4

IG(Y, temp) = 0.2116

Step 3: Repeat for each level (sad, I know)

Temperature



Entropy of "Sunny" node = $-((\frac{1}{3})\log_2\left(\frac{1}{3}\right) + (\frac{2}{3})\log_2(\frac{2}{3})) = 0.9183$

Entropy of its children = 0

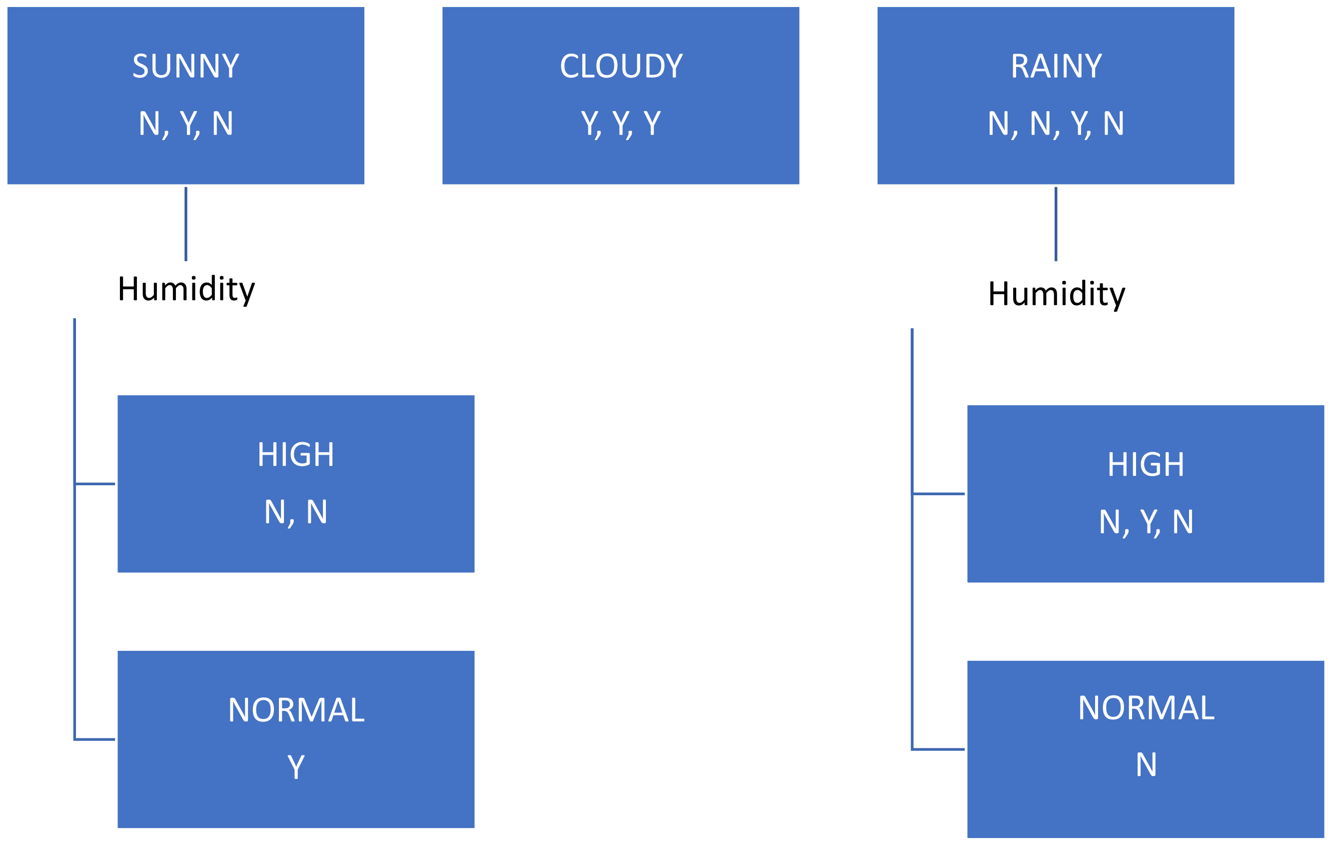
IG = 0.9183

Entropy of "Rainy" node = $-(\left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)\log_2(\frac{3}{4})) = \ 0.8113$

Entropy of children = $-(\frac{3}{4})(\left(\frac{1}{3}\right)\log_2\left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)\log_2(\frac{2}{3})) + \ 0 = \ 0.6887$

IG = 0.1226

Humidity



Entropy of "Sunny" node = $-((\frac{1}{3})\log_2\left(\frac{1}{3}\right) + (\frac{2}{3})\log_2(\frac{2}{3})) = 0.9183$

Entropy of its children = 0

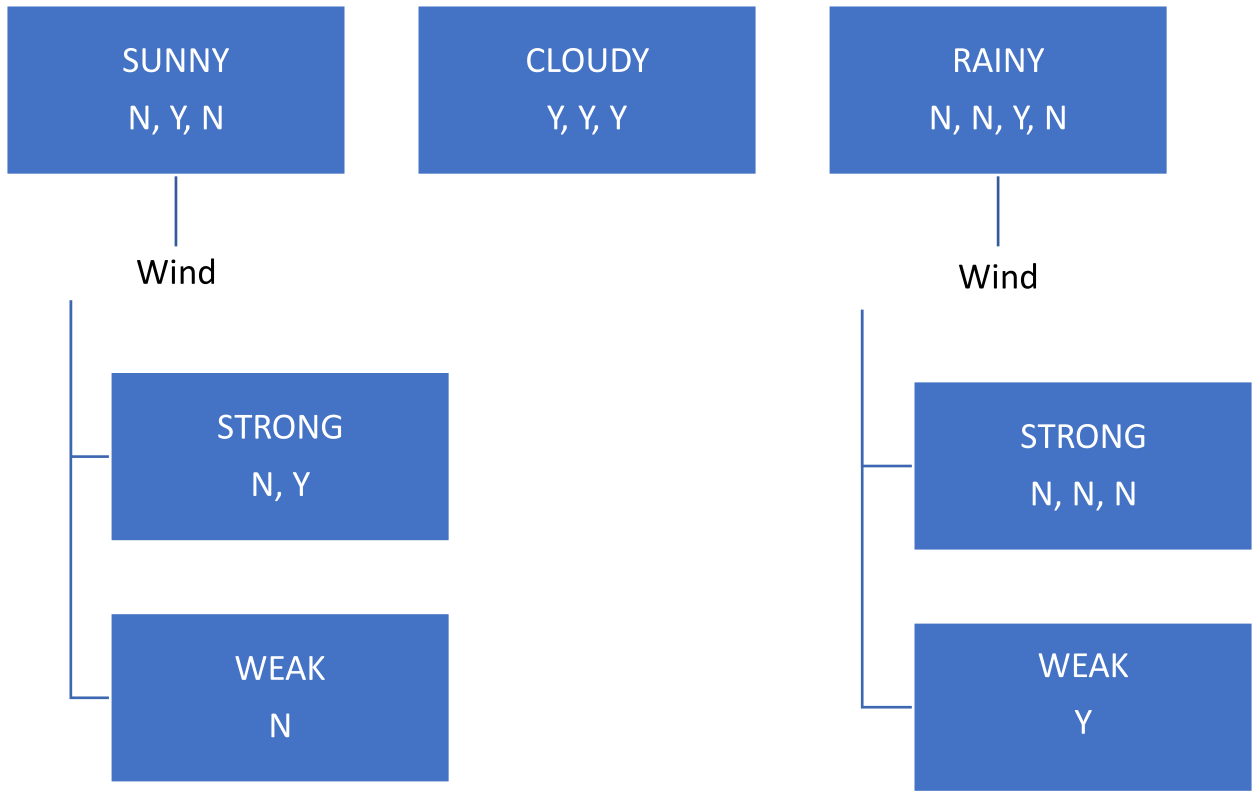
IG = 0.9183

Entropy of "Rainy" node = $-(\left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right) + \left(\frac{3}{4}\right)\log_2(\frac{3}{4})) = 0.8113$

Entropy of children = $-(\frac{3}{4})(\left(\frac{1}{3}\right)\log_2\left(\frac{1}{3}\right) + \left(\frac{2}{3}\right)\log_2(\frac{2}{3})) + 0 = 0.6887$

IG = 0.1226

Wind



Entropy of "Sunny" node = $-((\frac{1}{3})\log_2\left(\frac{1}{3}\right)+(\frac{2}{3})\log_2(\frac{2}{3}))\ =0.9183$

Entropy of its children = $-(\frac{2}{3})(\left(\frac{1}{2}\right)\log_2\left(\frac{1}{2}\right)+\left(\frac{1}{2}\right)\log_2(\frac{1}{2}))+0$ = 0.6667

IG = 0.2516

Entropy of "Rainy" node = $-(\left(\frac{1}{4}\right)\log_2\left(\frac{1}{4}\right)+\left(\frac{3}{4}\right)\log_2(\frac{3}{4}))\ =\ \ 0.8113$

Entropy of children = 0

IG = 0.8113

Step 4: Choose feature for each node to split on!

“Sunny node”:

IG(Y, weather) = IG(humidity) = 0.9183

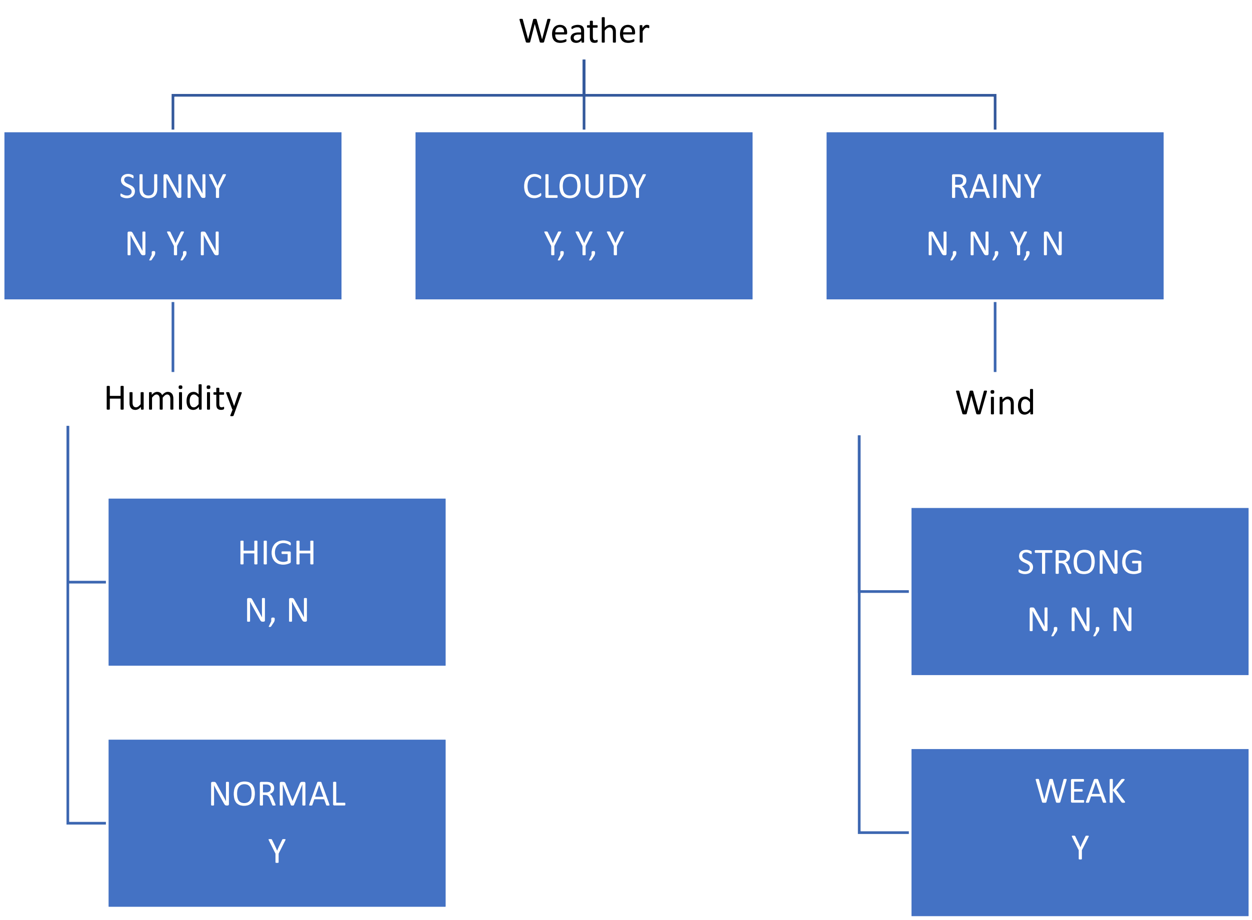
IG(Y, wind) = 0.2516

“Rainy node”:

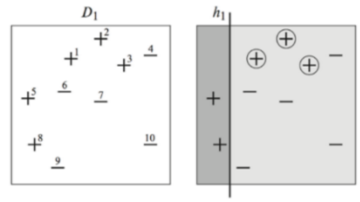
IG(Y, weather) = IG(Y, humidity) = 0.1226

IG(Y, wind) = 0.8113

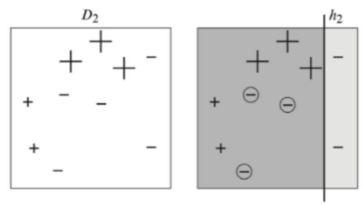
Final Tree!



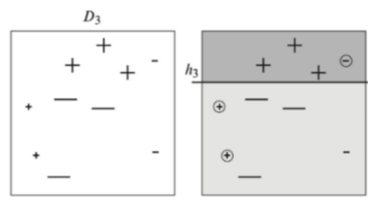
Boosting



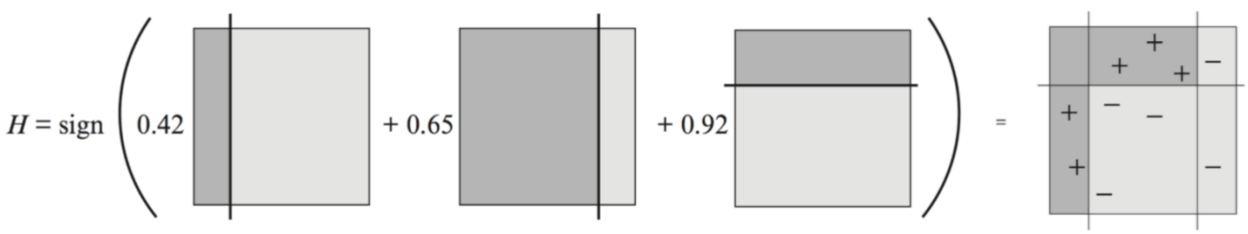
	1	2	3	4	5	6	7	8	9	10	
$D_1(i)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	$\epsilon_1 = 0.30, \alpha_1 \approx 0.42$
$e^{-\alpha_1 y_i h_1(x_i)}$	1.53	1.53	1.53	0.65	0.65	0.65	0.65	0.65	0.65	0.65	
$D_1(i)\, e^{-\alpha_1 y_i h_1(x_i)}$	0.15	0.15	0.15	0.07	0.07	0.07	0.07	0.07	0.07	0.07	$Z_1 \approx 0.92$



$D_2(i)$	0.17	0.17	0.17	0.07	0.07	0.07	0.07	0.07	0.07	0.07	$\epsilon_2 \approx 0.21, \alpha_2 \approx 0.65$
$e^{-\alpha_2 y_i h_2(x_i)}$	0.52	0.52	0.52	0.52	0.52	1.91	1.91	0.52	1.91	0.52	
$D_2(i)\, e^{-\alpha_2 y_i h_2(x_i)}$	0.09	0.09	0.09	0.04	0.04	0.14	0.14	0.04	0.14	0.04	$Z_2 \approx 0.82$



$D_3(i)$	0.11	0.11	0.11	0.05	0.05	0.17	0.17	0.05	0.17	0.05	$\epsilon_3 \approx 0.14, \alpha_3 \approx 0.92$
$e^{-\alpha_3 y_i h_3(x_i)}$	0.40	0.40	0.40	2.52	2.52	0.40	0.40	2.52	0.40	0.40	
$D_3(i)\, e^{-\alpha_3 y_i h_3(x_i)}$	0.04	0.04	0.04	0.11	0.11	0.07	0.07	0.11	0.07	0.02	$Z_3 \approx 0.69$



(https://www.ccs.neu.edu/home/vip/teach/MLcourse/4_boosting/slides/boosting.pdf)