Decision Tree
- A supervised ML algorithm, used for both regression and classification, a probability based approach
- A non linear technique, it can yield good performance even if features do not have good correlation with target
- A white box model - can be explained to business
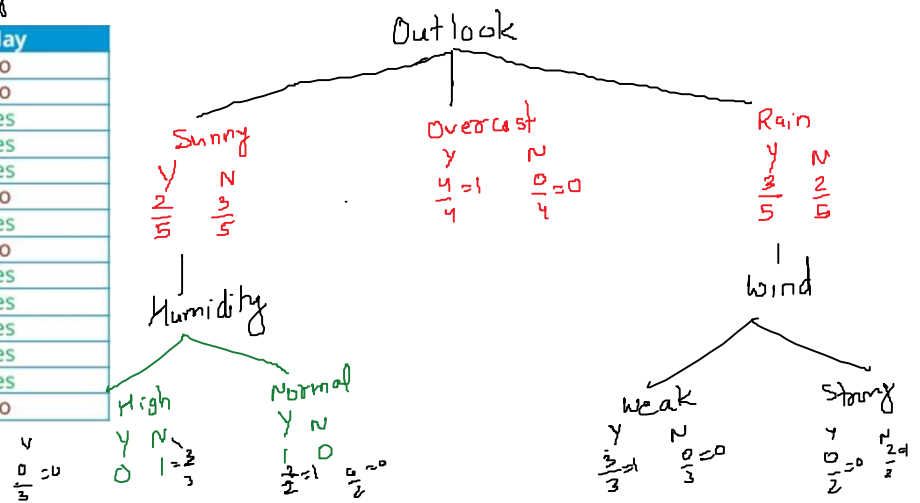- Does not require feature scaling

A   B
0.95   0.05
low entropy

Entropy: amount of uncertainty in any event.

$P_Y = \frac{9}{14}$ $P_N = \frac{5}{14}$

A   B
0.45   0.55
high entropy

	x_1	x_2	x_3	y
Day	**Outlook**	**Humidity**	**Wind**	**Play**
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

Outlook

Sunny
Y N
$\frac{2}{5}$ $\frac{3}{5}$

Overcast
Y N
$\frac{4}{4}=1$ $\frac{0}{4}=0$

Rain
Y N
$\frac{3}{5}$ $\frac{2}{5}$

Humidity

High
Y N
$\frac{0}{3}=0$ $\frac{1}{3}=\frac{3}{3}$?

Normal
Y N
$\frac{2}{2}=1$ $\frac{0}{2}=0$

Wind

weak
Y N
$\frac{3}{3}=1$ $\frac{0}{3}=0$

Strong
Y N
$\frac{0}{2}=0$ $\frac{2}{2}=1$

train ● ●
test ● ●

$x_2 < 0.5$ $x_2 > 0.5$
B   R

$x_1 > 0.5$ $x_1 < 0.5$ $x_1 < 0.3$ $x_1 > 0.3$
R   B   R   B

$x_2 < 0.4$ $x_2 > 0.4$ $x_1 > 0.3$ $x_1 < 0.5$ $x_1 < 0.6$ $x_1 > 0.6$
B   B   R

$x_1 < 0.1$ $x_1 > 0.1$ $x_1 < 0.6$ $x_1 < 0.6$
R   B