AdaBoost Example

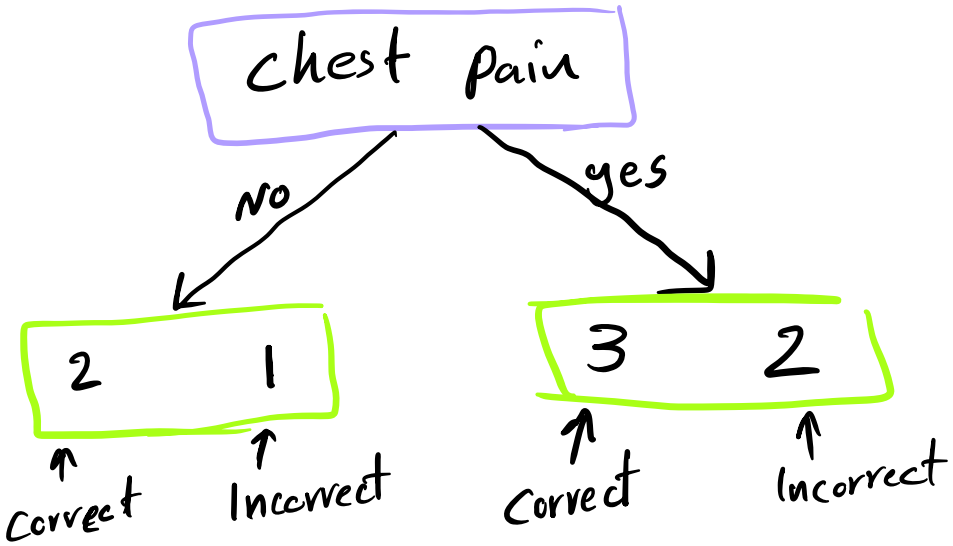
Chest Pain	Blocked Arteries	Patient Weight	Heart Disease	sample weight
Yes	Yes	205	Yes	$\frac{1}{8}$
No	Yes	180	Yes	$\frac{1}{8}$
Yes	No	210	Yes	$\frac{1}{8}$
Yes	Yes	167	Yes	$\frac{1}{8}$
No	Yes	156	No	$\frac{1}{8}$
No	Yes	125	No	$\frac{1}{8}$
Yes	No	168	No	$\frac{1}{8}$
Yes	Yes	172	No	$\frac{1}{8}$

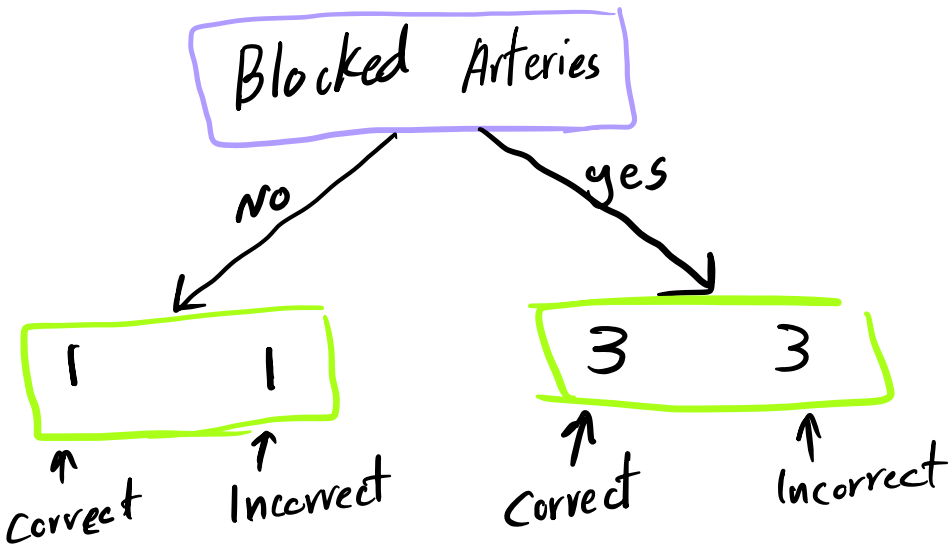
First give each sample a weight.

All samples get same weight

Note: The sum of all weights should be 1

Now we creat our first stump





---

geni For weight =>

$L \Rightarrow 1 - \left(\frac{4}{5}\right)^2 - \left(\frac{1}{5}\right)^2$

$= 1 - \frac{17}{25} \Rightarrow \frac{8}{25}$

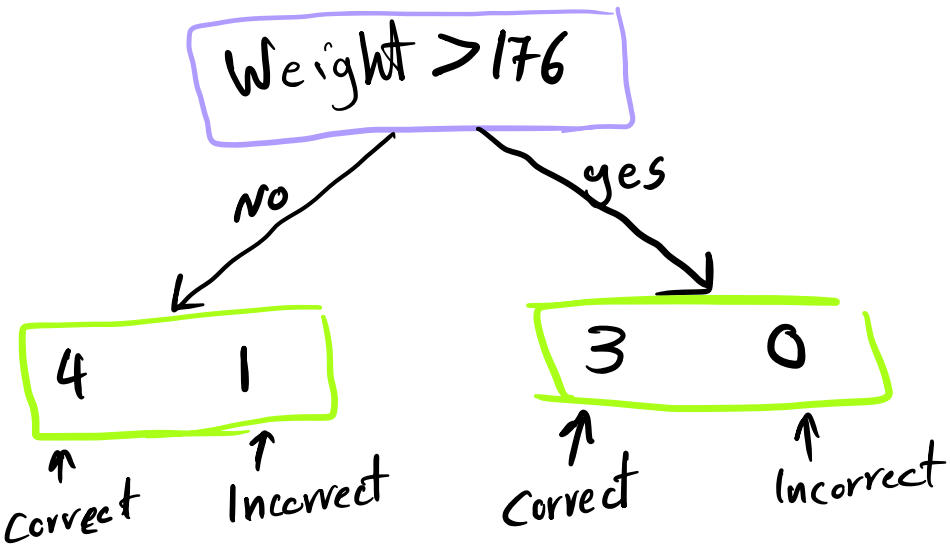
$R = 1 - \left(\frac{3}{3}\right)^2 - \left(\frac{0}{3}\right)^2$

$R = 1 - 1 \Rightarrow 0$

total

$\frac{8}{25} \cdot \frac{5}{8} + 0 \cdot \frac{3}{8}$

$\Rightarrow \frac{1}{5} \Rightarrow 0.2$



Now calculate gini Index for each stump.

Chest Pain stump gini Index = 0,47

Blocked Arteries gini Index = 0,5

Weight gini Index = 0,2

We choose the lowest gini Index that have less error. Weight

Now weight is first stump.

We should calculat Amount of say for first stump.

$$\text{Amount of say} = \frac{1}{2}\log\left(\frac{1-\text{total error}}{\text{total error}}\right)$$

$$= \frac{1}{2}\log\left(\frac{1-\frac{1}{8}}{\frac{1}{8}}\right)$$

$$= \frac{1}{2}\log\left(\frac{\frac{7}{8}}{\frac{1}{8}}\right)$$

$$= \frac{1}{2}\log(7)$$

$$= 0.97$$

Now we should update the weights for each sample.

New sample weight = sample weight $\times e^{\pm \text{Amount to scale}}$

Postive for miss classifide samples and Negtive for the others

start with miss classified

$\frac{1}{8} \cdot e^{0.97} \Rightarrow \frac{1}{8} \times 2.64 = 0,33$ miss classified

$\frac{1}{8} \cdot e^{-0,97} \Rightarrow \frac{1}{8} \times 0,38 = 0,05$ correct

Chest Pain	Blocked Arteries	Patient Weight	Heart Disease
Yes	Yes	205	Yes
No	Yes	180	Yes
Yes	No	210	Yes
Yes	Yes	167	Yes
No	Yes	156	No
No	Yes	125	No
Yes	No	168	No
Yes	Yes	172	No

Sample weight → New weights → Normalized

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

$\frac{1}{8} \Rightarrow 0,33 \Rightarrow 0,49$

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

$\frac{1}{8} \Rightarrow 0,05 \Rightarrow 0,07$

Now we should Normalize the weights so that they will add up to 1

The total weight is equall 0,68 so we divide by the total to normalize the data

Now we creat new dataset based on previous one

Chest Pain	Blocked Arteries	Patient Weight	Heart Disease
No	Yes	156	No
Yes	Yes	167	Yes
No	Yes	125	No
Yes	Yes	167	Yes
Yes	Yes	167	Yes
Yes	Yes	172	No
Yes	Yes	205	Yes
Yes	Yes	167	Yes

It Reapeted 4 times because it has high weight

Now repeate till end of features.