AdaBoost Tutorial

AdaBoost Algorithm

Algorithm 1 Discrete AdaBoost

Input: Training data D, number of iterations T, and an initial distribution $w_0(i)$ over the examples.
For $t = 1, ..., T$:

- Train a weak classifier $h_t = \arg\min_h \sum_{i=1}^{n} w_{t-1}(i)\mathbf{1}(y_i \neq h(x_i))$.
- Calculate the error of $h_t : \epsilon_t = \sum_{i=1}^{n} w_{t-1}(i)\mathbf{1}(y_i \neq h_t(x_i))$.
- Set $\alpha_t = \frac{1}{2}\log\frac{1-\epsilon_t}{\epsilon_t}$.
- Set $w_{t+1}(i) = \begin{cases} \frac{w_t(i)}{2(1-\epsilon_t)} & \text{if } y_i = h_{t+1}(x_i) \\ \frac{w_t(i)}{2\epsilon_t} & \text{if } y_i \neq h_{t+1}(x_i) \end{cases}$

Output: Strong classifier $H(x) = \sum_{k=1}^{T} \alpha_k h_k(x)$.

AdaBoost Algorithm

Initialize weights to all training points

$$w = \frac{1}{N}$$

Calculate error rate for each weak classifier

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Classifier with the lowest error rate

Now compute Voting Power for the classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

Append the classifier in the ensemble classifier and heck if the classifier is good enough?

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

Update weights of each point where the previous classifier went wrong

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} & \text{if point classified correctly} \\ \frac{w_{old}}{2\epsilon} & \text{if point classified wrongly} \end{cases}$$

$$w = \frac{1}{N}$$

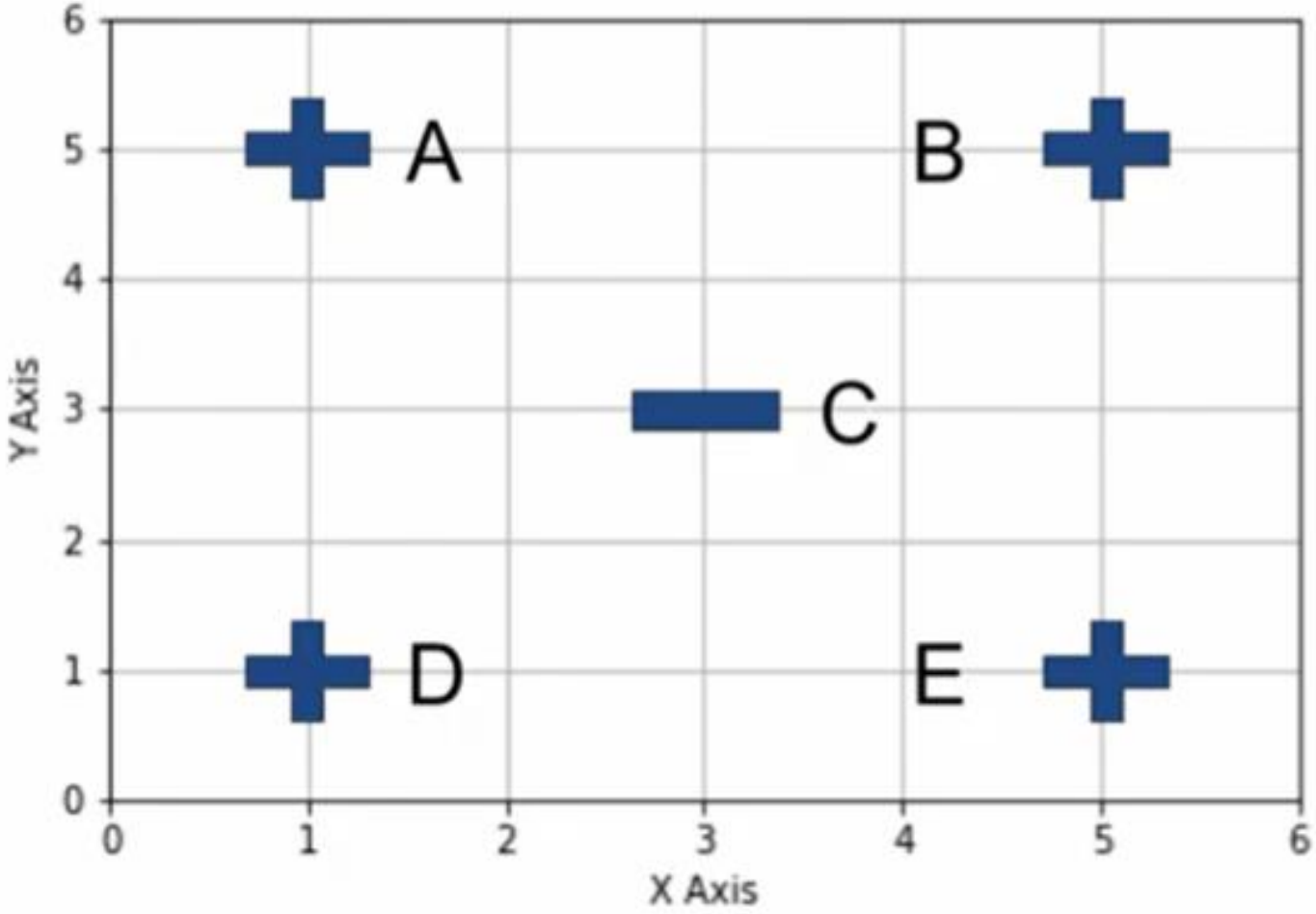
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



- Let's start with 5 points.
- 4 points belong to class 1 & are denoted by plus sign.
- 1 point belongs to class 0 and is denoted by minus sign.

$$w = \frac{1}{N}$$

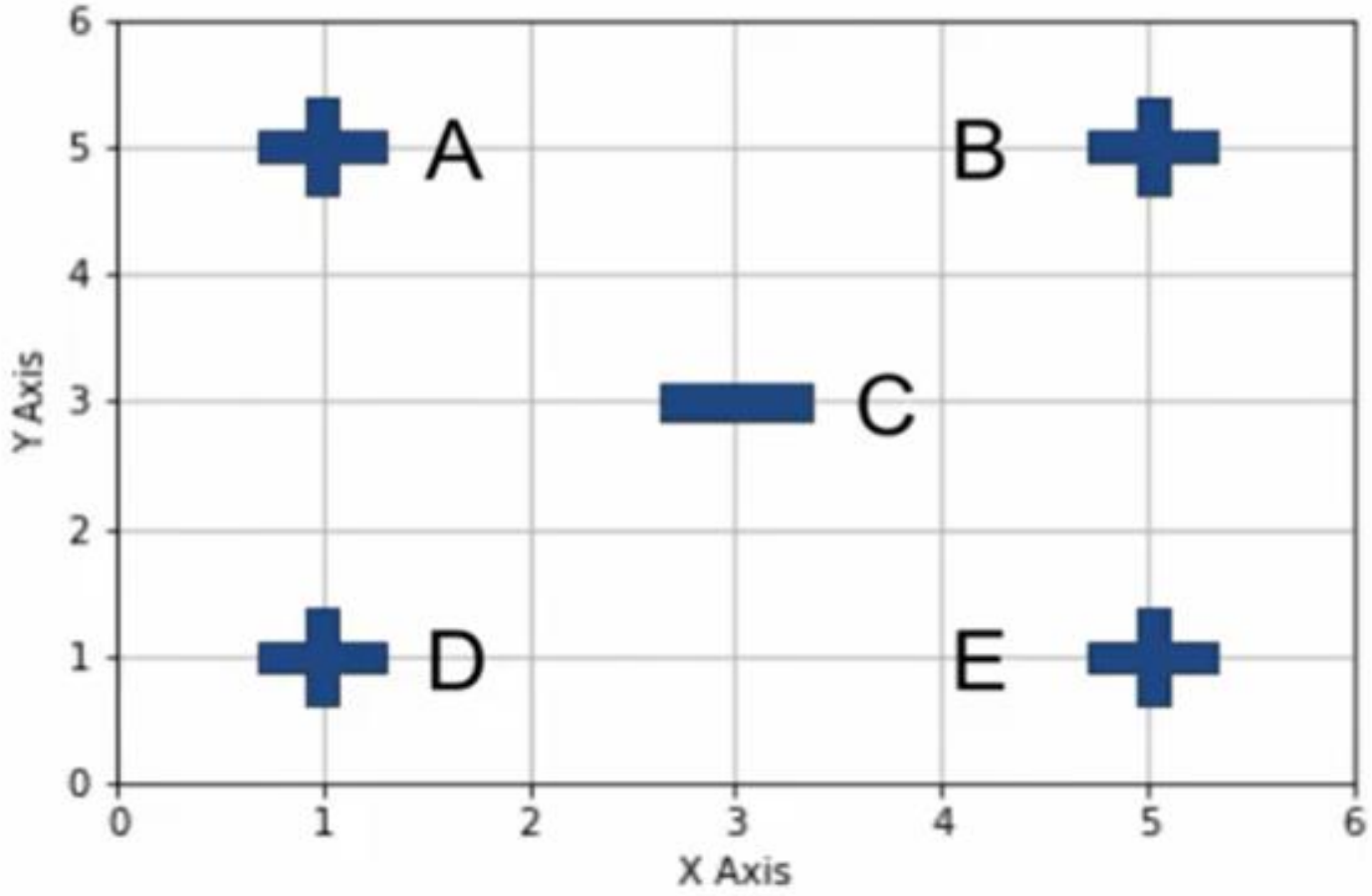
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

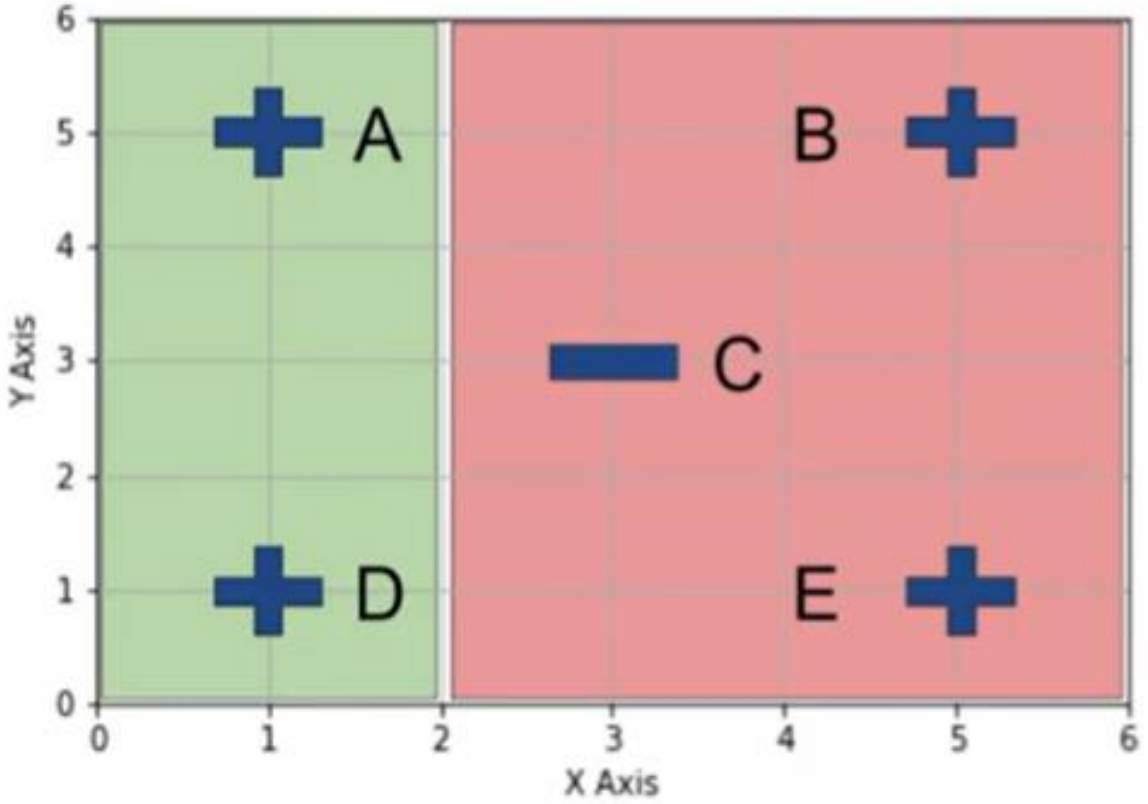
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	
X < 4		
X < 6		
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

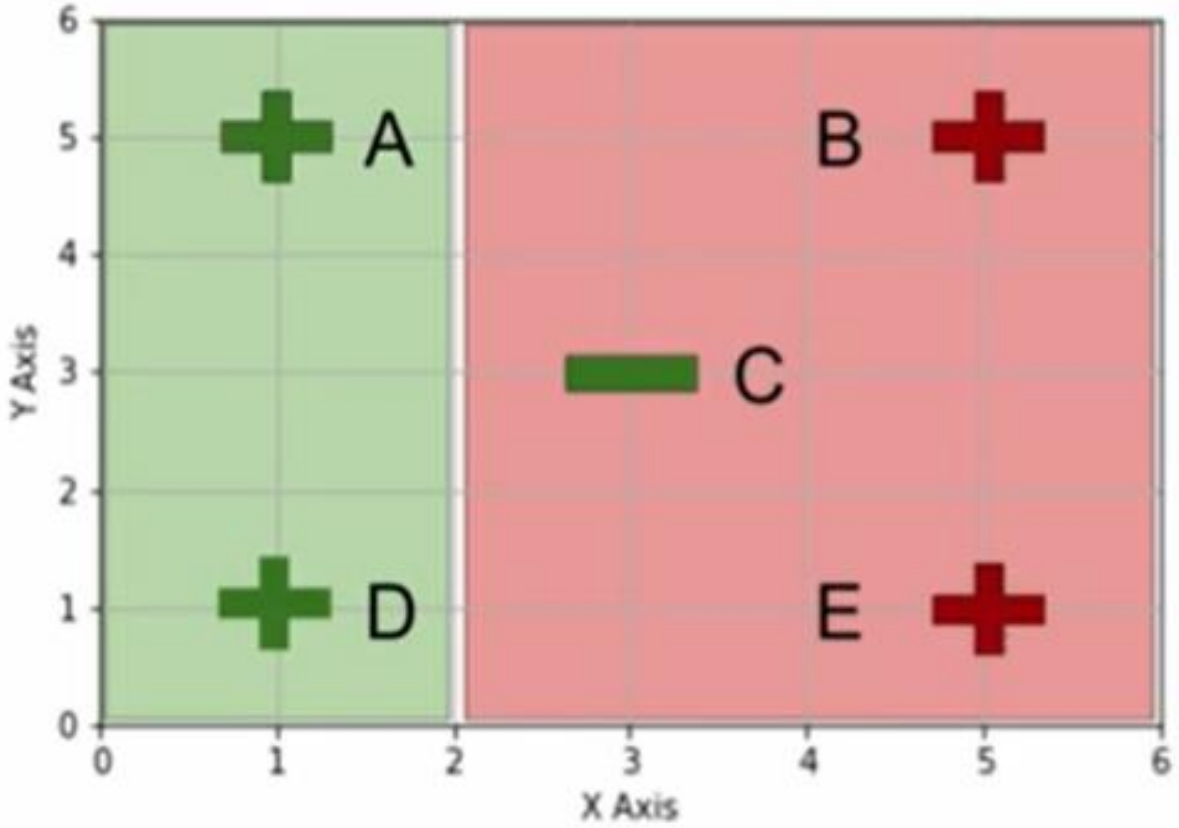
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/5
X < 4		
X < 6		
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

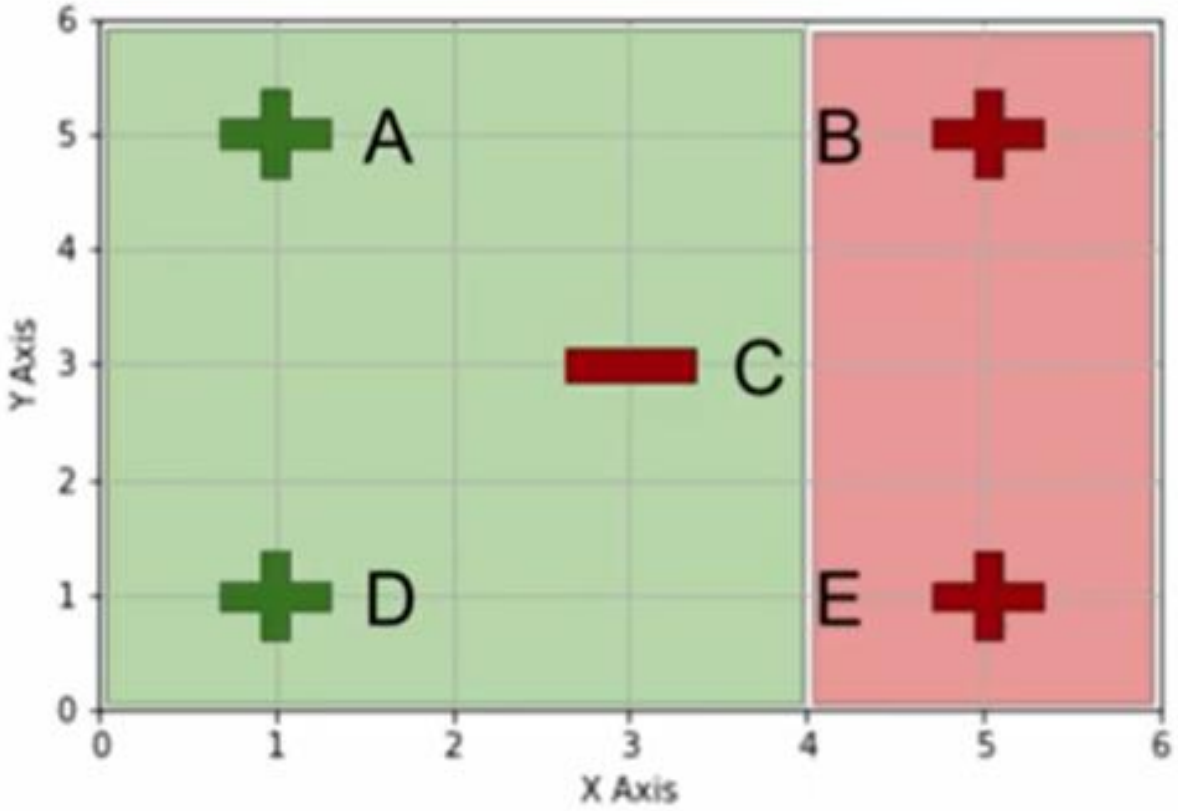
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/5
X < 4	B, C & E	3/5
X < 6		
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

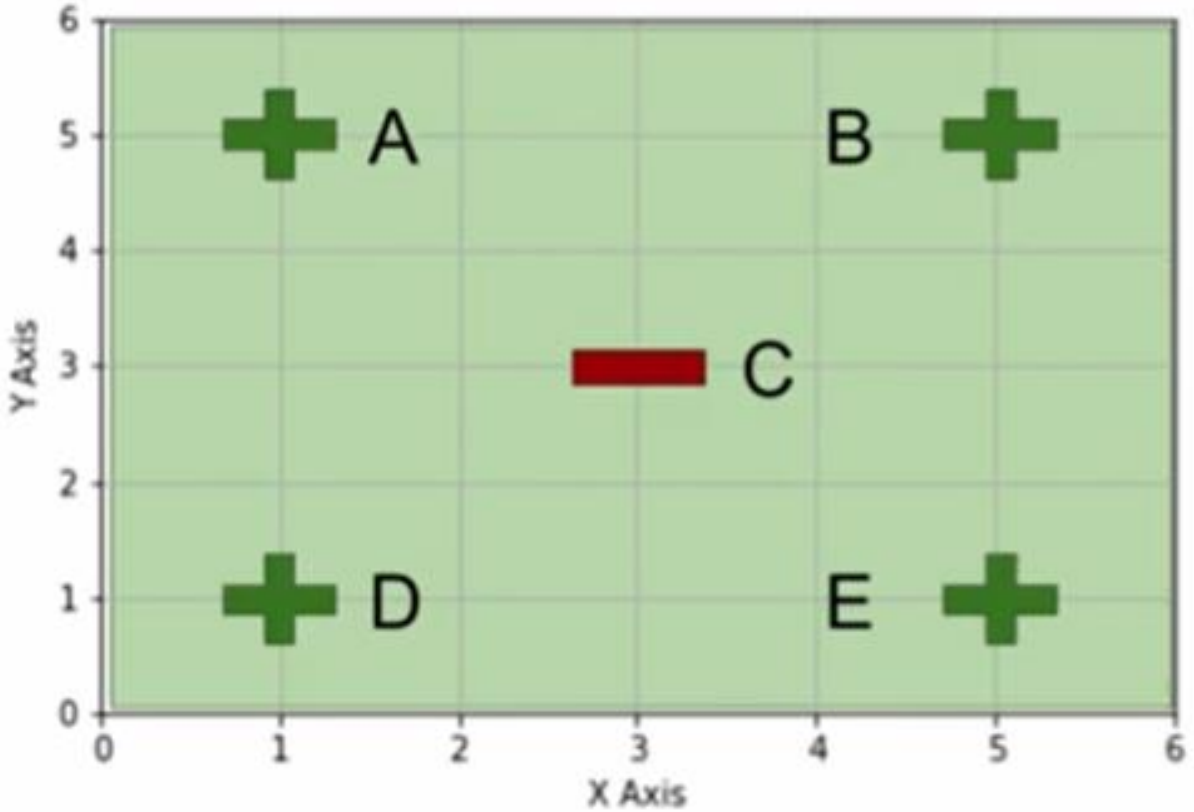
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/5
X < 4	B, C & E	3/5
X < 6	C	1/5
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

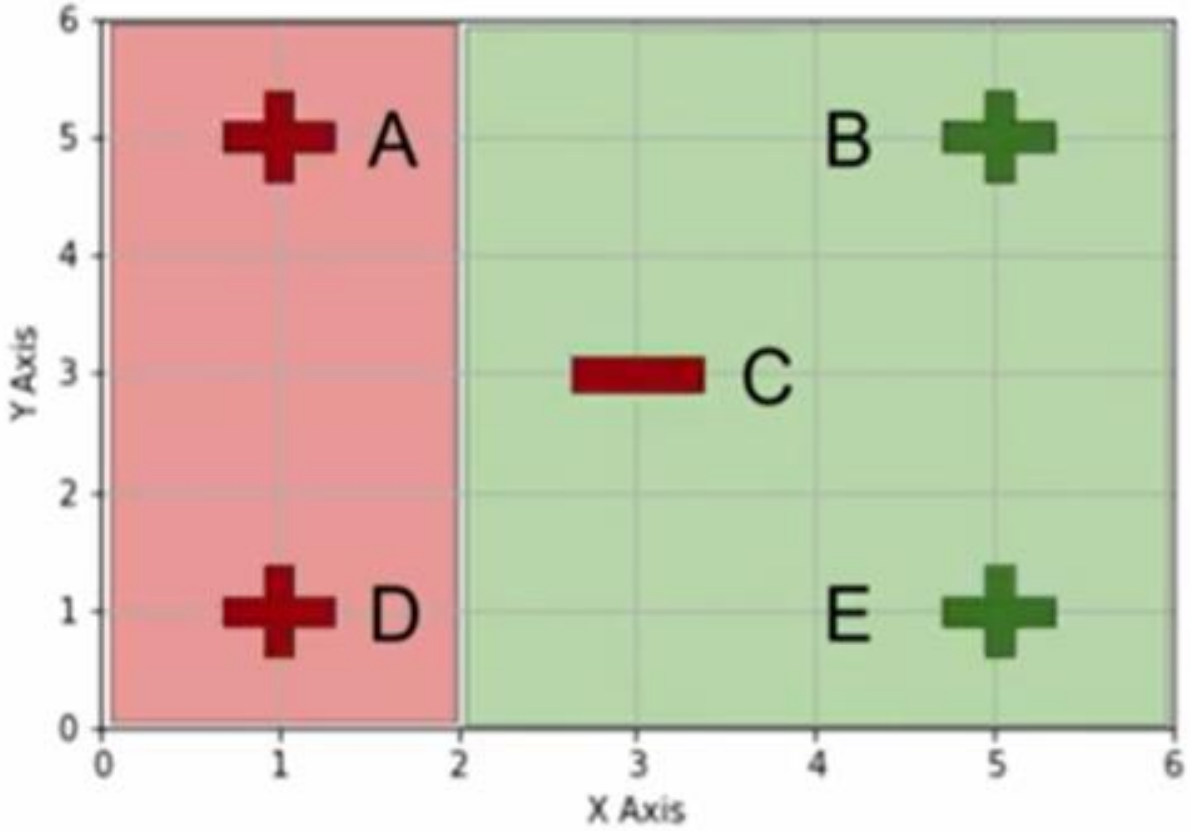
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/5
X < 4	B, C & E	3/5
X < 6	C	1/5
X > 2	A, D & C	3/5
X > 4		
X > 6		

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

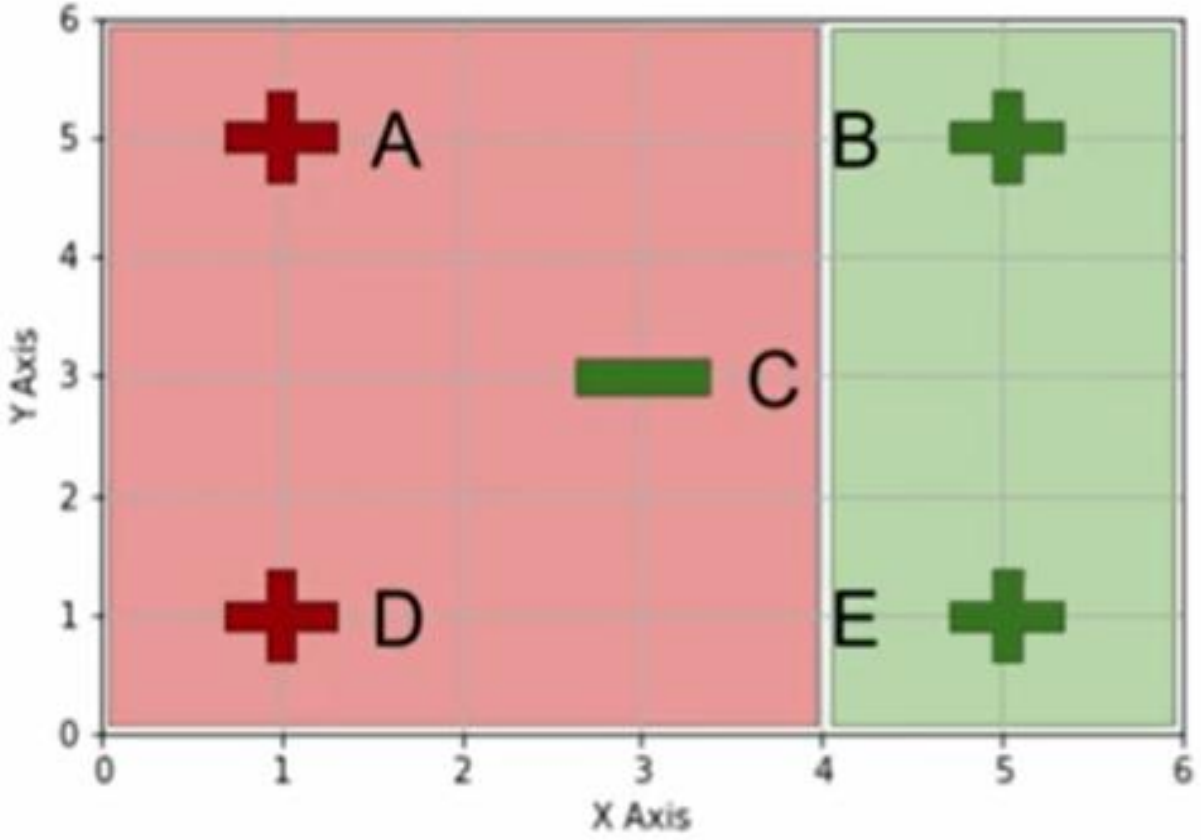
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2}\log\frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/5
X < 4	B, C & E	3/5
X < 6	C	1/5
X > 2	A, D & C	3/5
X > 4	A, D	2/5
X > 6		

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

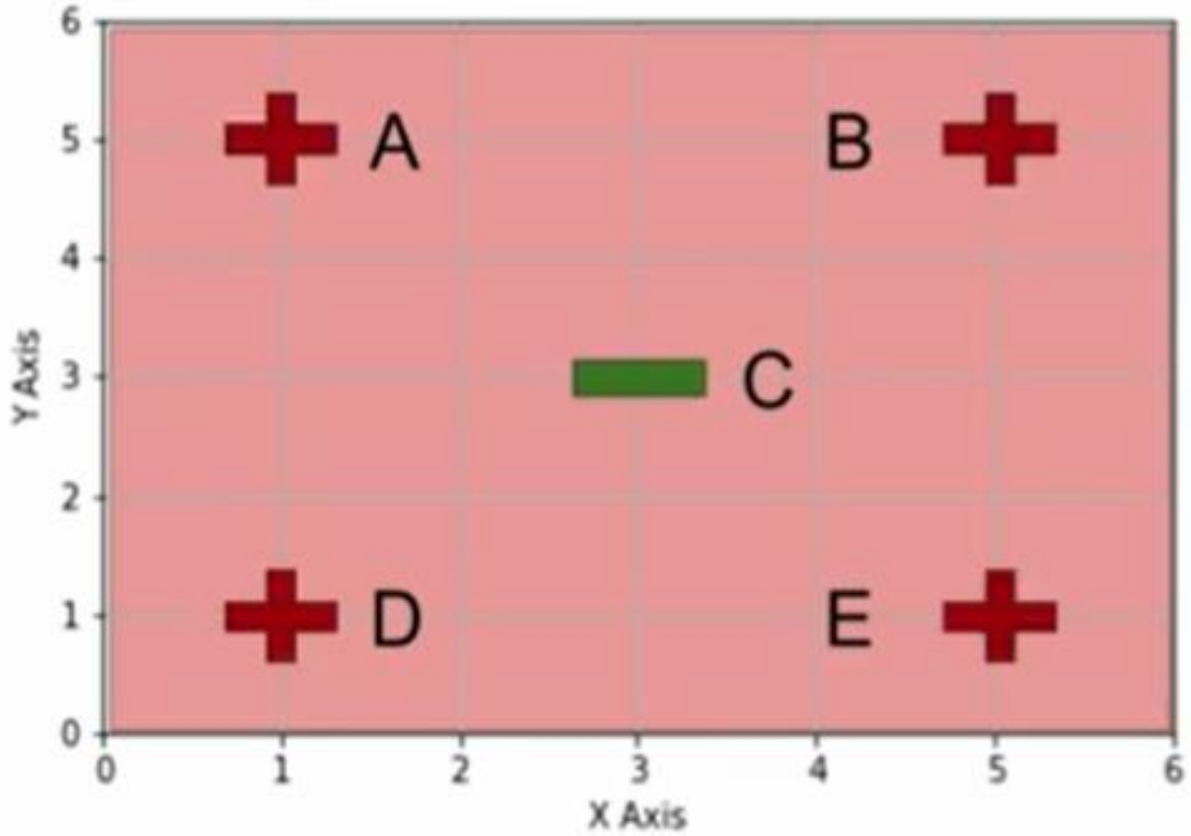
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/5
X < 4	B, C & E	3/5
X < 6	C	1/5
X > 2	A, D & C	3/5
X > 4	A, D	2/5
X > 6	A, B, D, E	4/5

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

C	Wrong	Error
X < 2	B & E	2/5
X < 4	B, C & E	3/5
X < 6	C	1/5
X > 2	A, D & C	3/5
X > 4	A, D	2/5
X > 6	A, B, D, E	4/5

- Going through the error rates of each classifier, we find that X< 6 is the best performing classifier.
- Now, that we have decided on the classifier, let's now find the voting power of the classifier.

Points	Weight
Wa	1/5
Wb	1/5
Wc	1/5
Wd	1/5
We	1/5

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

$$\epsilon = 1/5$$

$$\alpha = \frac{1}{2} \log \frac{1 - \frac{1}{5}}{\frac{1}{5}}$$

$$\alpha = \frac{1}{2} \log 4$$

$$h(x) = \frac{1}{2} \log 4 * F(x < 6)$$

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} & \text{if point classified correctly} \\ \frac{w_{old}}{2\epsilon} & \text{if point classified wrongly} \end{cases}$$

- Points A, B, D and E are correctly classifier.
- Plugging in the value of error = ⅕ and initial weight of ⅕ in the 1st equation.
- We get new weight = ⅛

Points	Weight
Wa	1/8
Wb	1/8
Wc	
Wd	1/8
We	1/8

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} & \text{if point classified correctly} \\ \frac{w_{old}}{2\epsilon} & \text{if point classified wrongly} \end{cases}$$

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

- Points C is incorrectly classifier.
- Plugging in the value of error = ⅕ and initial weight of ⅕ in the 1st equation.
- We get new weight = ½

$$w = \frac{1}{N}$$

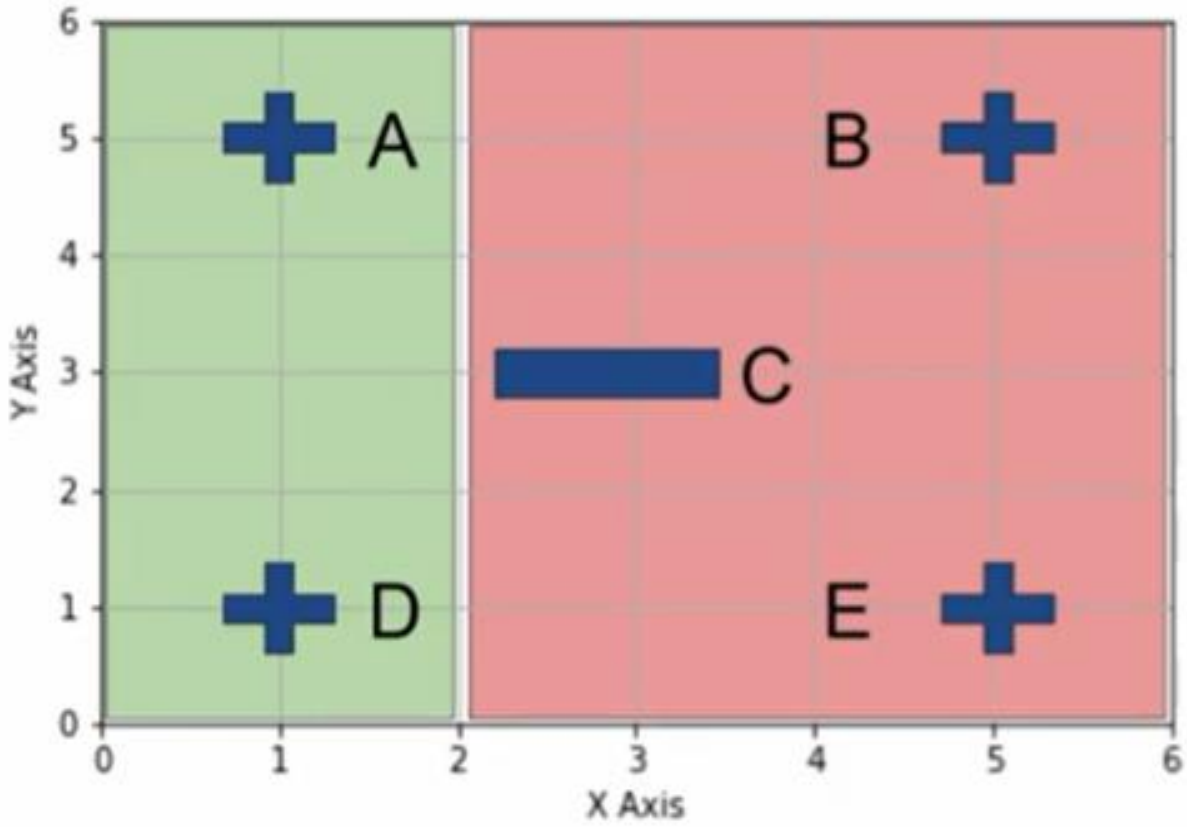
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	
X < 4		
X < 6		
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

Bhavesh

$$w = \frac{1}{N}$$

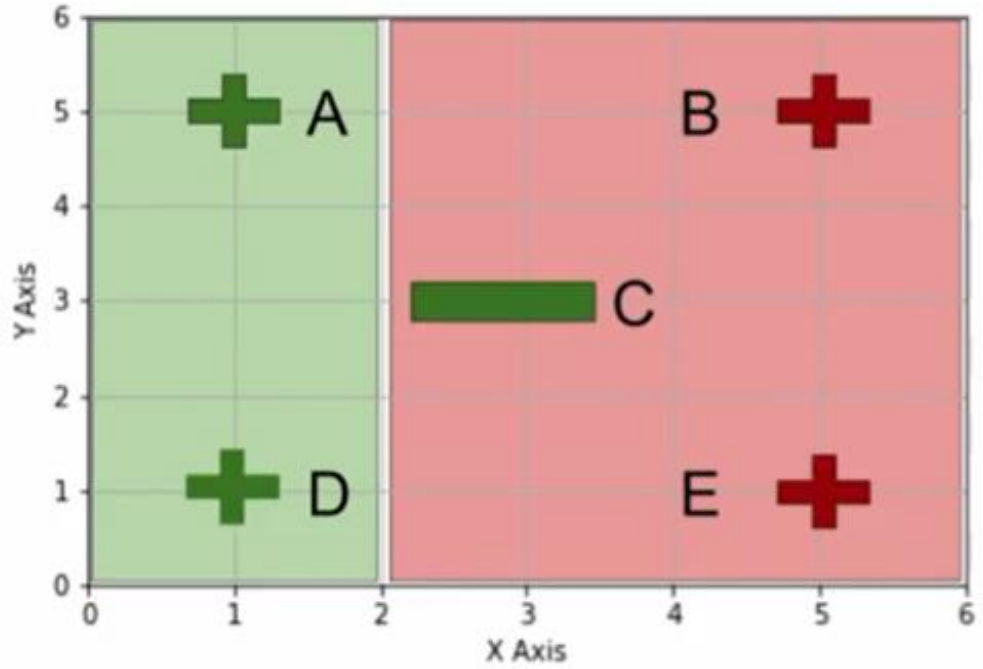
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/8
X < 4		
X < 6		
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

Bhavesh

$$w = \frac{1}{N}$$

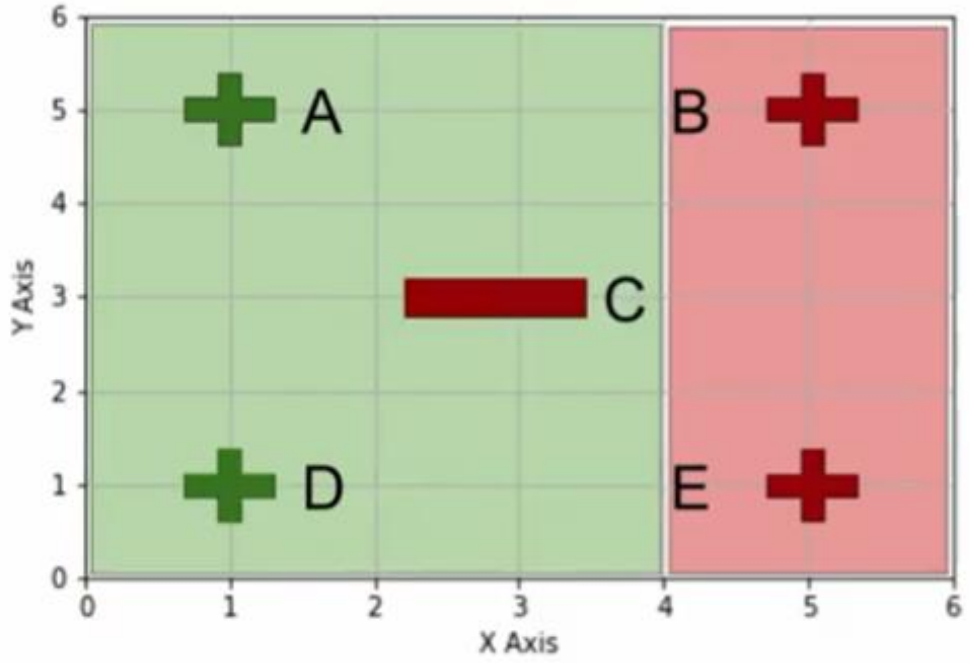
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/8
X < 4	B, C & E	6/8
X < 6		
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

Bhavesh

$$w = \frac{1}{N}$$

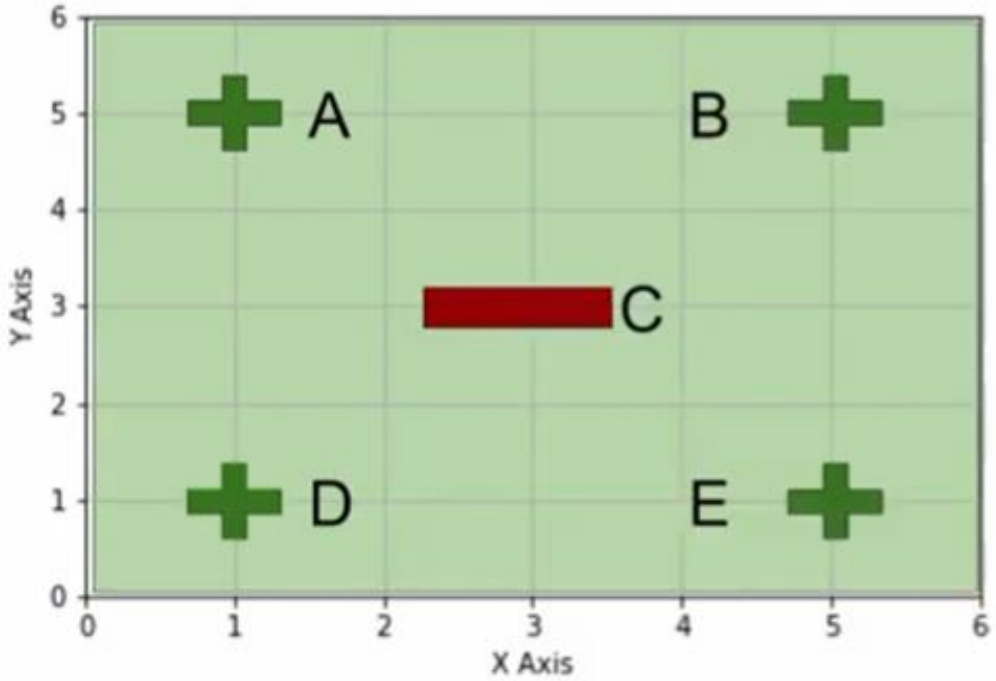
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/8
X < 4	B, C & E	6/8
X < 6	C	4/8
X > 2		
X > 4		
X > 6		

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

$$w = \frac{1}{N}$$

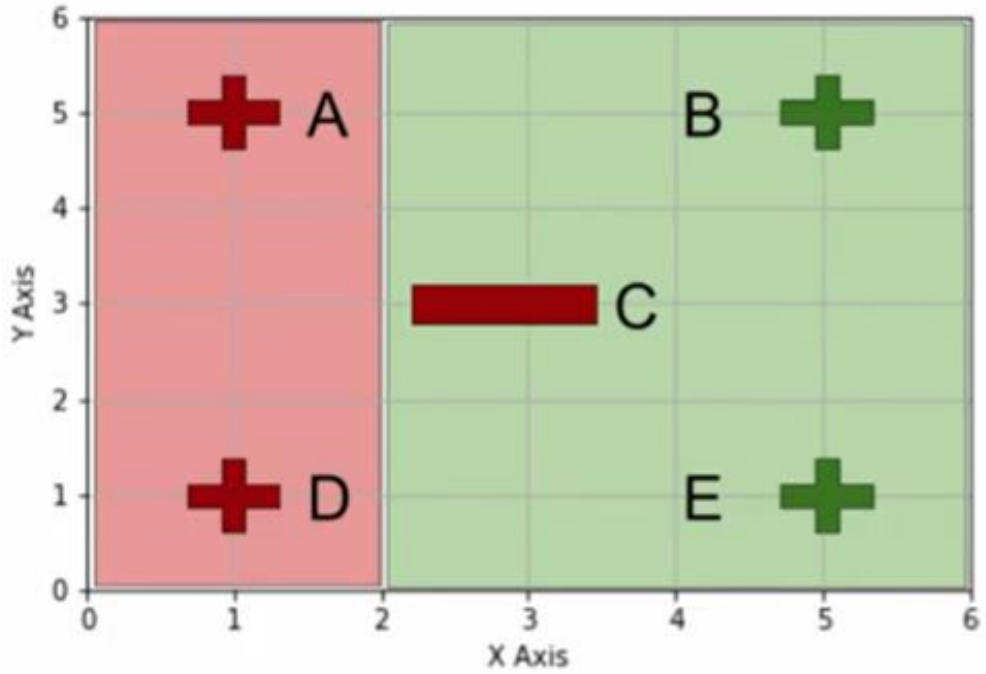
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{\text{old}}}{2(1-\epsilon)} \\ \frac{w_{\text{old}}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/8
X < 4	B, C & E	6/8
X < 6	C	4/8
X > 2	A, C & D	6/8
X > 4		
X > 6		

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

$$w = \frac{1}{N}$$

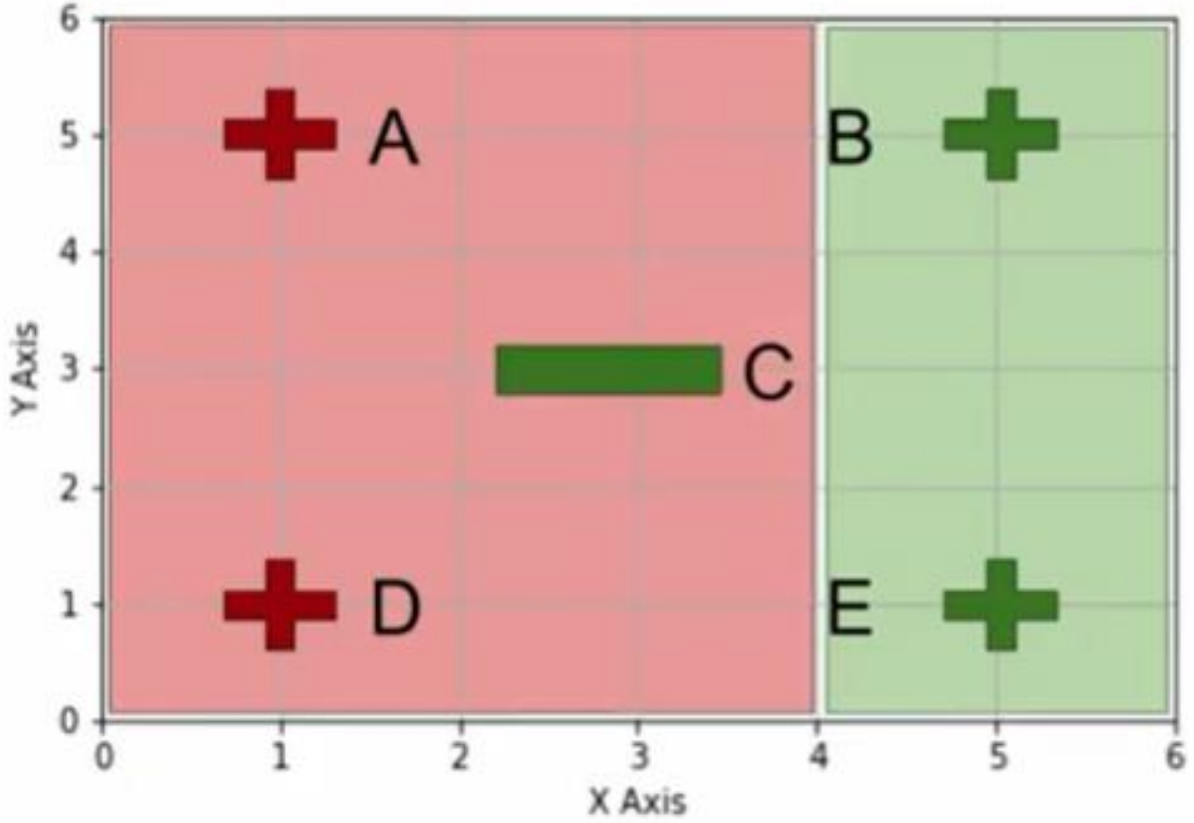
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/8
X < 4	B, C & E	6/8
X < 6	C	4/8
X > 2	A, C & D	6/8
X > 4	A, D	2/8
X > 6		

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

$$w = \frac{1}{N}$$

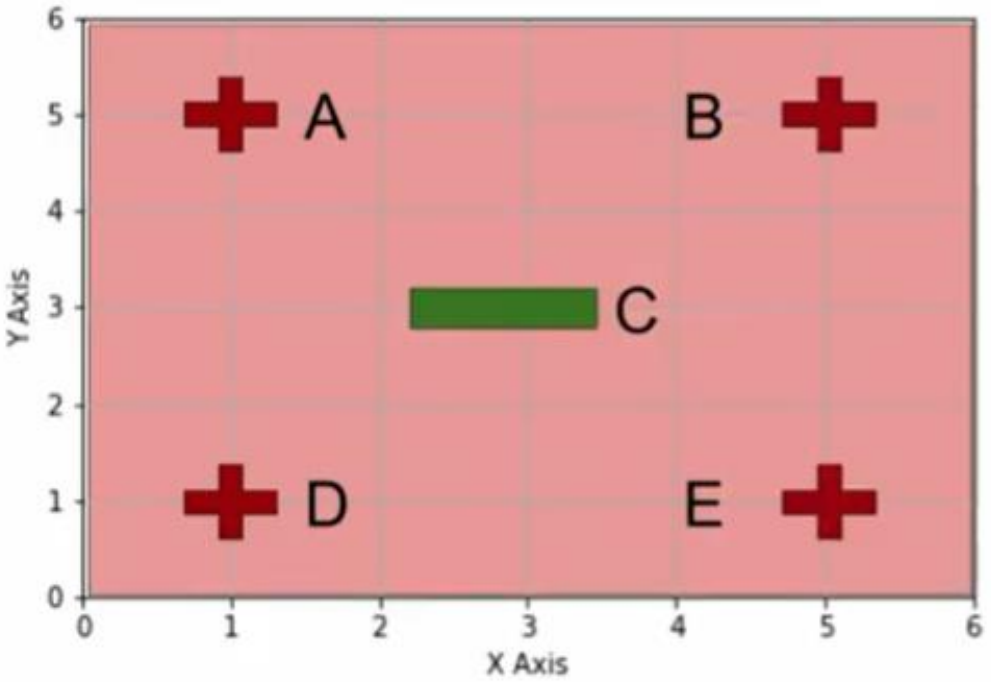
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	2/8
X < 4	B, C & E	6/8
X < 6	C	4/8
X > 2	A, C & D	6/8
X > 4	A, D	2/8
X > 6	A, B, D, E	4/8

Points	Weight
Wa	1/8
Wb	1/8
Wc	1/2
Wd	1/8
We	1/8

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

C	Wrong	Error
X < 2	B & E	2/8
X < 4	B, C & E	6/8
X < 6	C	4/8
X > 2	A, D & C	6/8
X > 4	A, D	2/8
X > 6	A, B, D, E	4/8

- Going through the error rates of each classifier, we find that there are 2 candidates X< 2 and X> 4 with minimum error rate.
- In this round, we will choose the 1st one i.e. X< 2 classifier.

Bhavesh

$w = \frac{1}{N}$

$\epsilon = \sum_{\text{wrong}} w_i$

Pick Lowest Error Rate Classifier

$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$

$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$

$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$

$$\epsilon = 2/8$$

$$\alpha = \frac{1}{2} \log \frac{1 - \frac{1}{4}}{\frac{1}{4}}$$

$$\alpha = \frac{1}{2} \log 3$$

$$h(x) = \frac{1}{2} \log 4 * F(x < 6) + \frac{1}{2} \log 3 * F(x < 2)$$

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} a_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} & \text{if point classified correctly} \\ \frac{w_{old}}{2\epsilon} & \text{if point classified wrongly} \end{cases}$$

- Points B and E are incorrectly classified.
- Plugging in the value of error = 2/8 and initial weights from previous round we get the following points weight table

Points	Weight
Wa	1/12
Wb	3/12
Wc	4/12
Wd	1/12
We	3/12

$$w = \frac{1}{N}$$

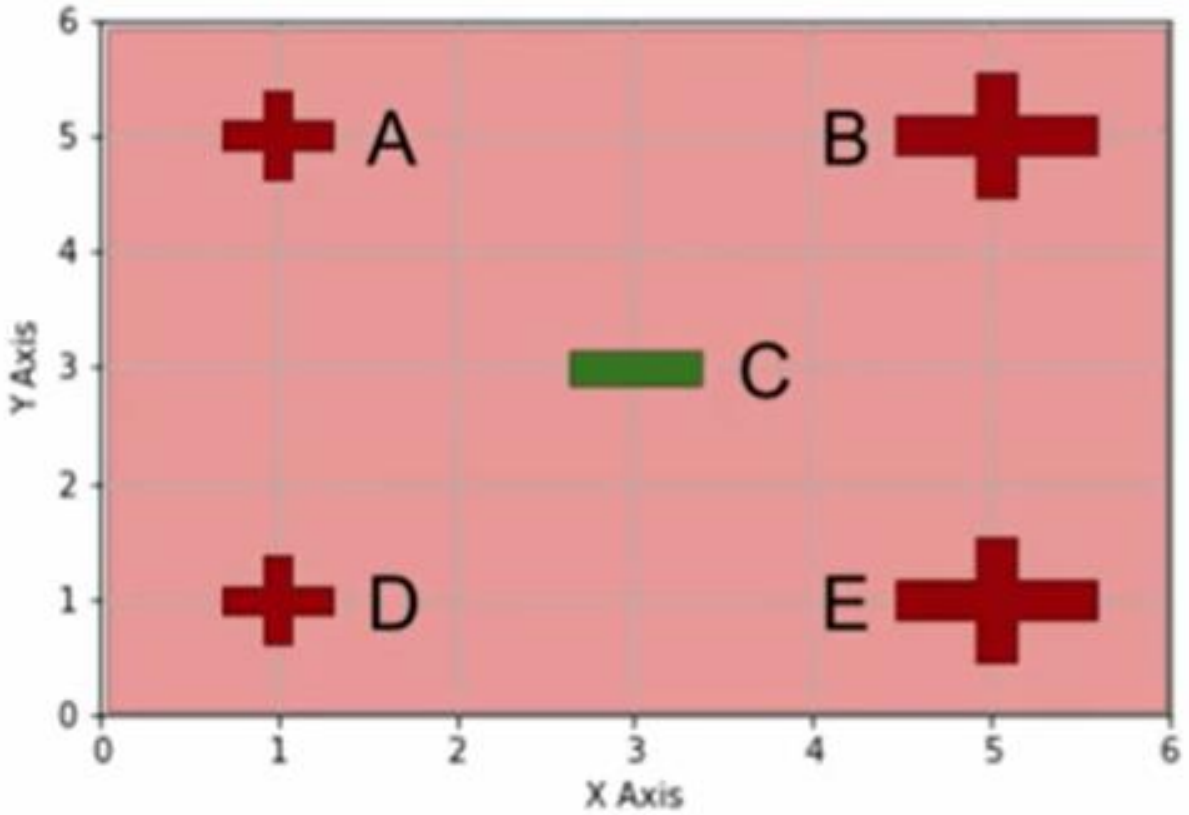
$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$



C	Wrong	Error
X < 2	B & E	1/2
X < 4	B, C & E	10/12
X < 6	C	4/12
X > 2	A, C & D	1/2
X > 4	A, D	2/12
X > 6	A, B, D, E	8/12

Points	Weight
Wa	1/12
Wb	3/12
Wc	4/12
Wd	1/12
We	3/12

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

C	Wrong	Error
X < 2	B & E	1/2
X < 4	B, C & E	10/12
X < 6	C	4/12
X > 2	A, D & C	1/2
X > 4	A, D	2/12
X > 6	A, B, D, E	8/12

- Going through the error rates of each classifier, we find that X> 4 is the classifier with minimum error rate.
- In this round, we will choose X> 4 classifier.

$$w = \frac{1}{N}$$

$$\epsilon = \sum_{\text{wrong}} w_i$$

Pick Lowest Error Rate Classifier

$$\alpha = \frac{1}{2} \log \frac{1-\epsilon}{\epsilon}$$

$$f(x_i) = \sum_{t=1}^{T} \alpha_t h_t(x_i)$$

$$w_{new} = \begin{cases} \frac{w_{old}}{2(1-\epsilon)} \\ \frac{w_{old}}{2\epsilon} \end{cases}$$

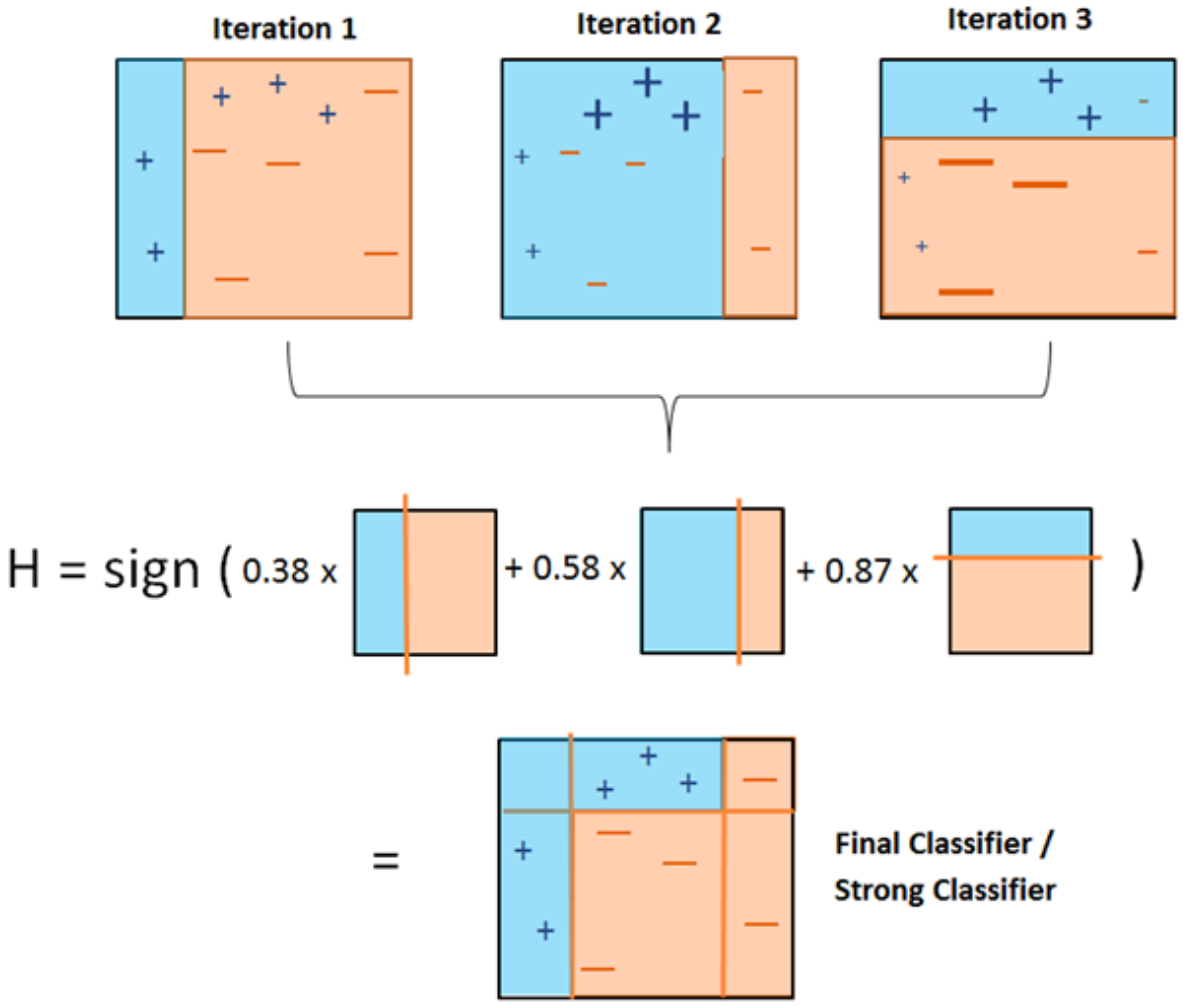
$$\epsilon = 1/6$$

$$\alpha = \frac{1}{2} \log \frac{1-\frac{1}{6}}{\frac{1}{6}}$$

$$\alpha = \frac{1}{2} \log 5$$

$$h(x) = \frac{1}{2} \log 4 * F(x < 6) + \frac{1}{2} \log 3 * F(x < 2) + \frac{1}{2} \log 5 * F(x > 4)$$

AdaBoost Classifier Working Principle with Decision Stump as a Base Classifier




Thank You