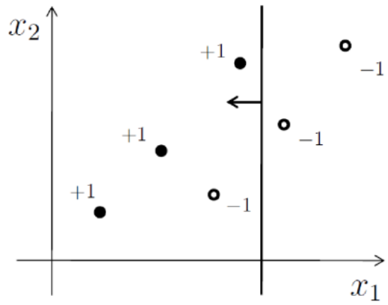
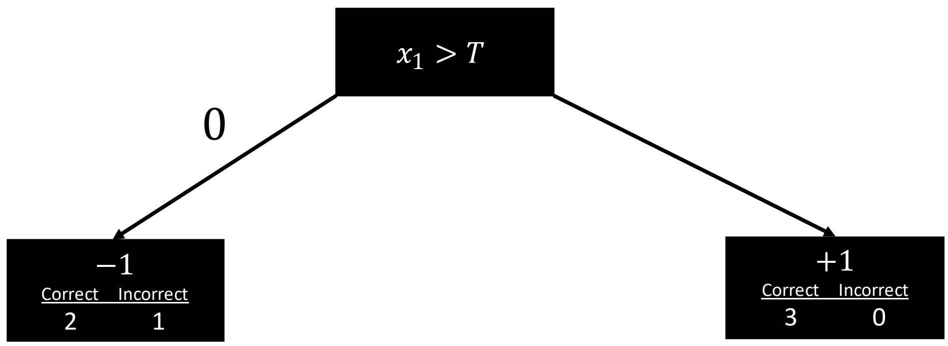
HW4 - Question 2 – AdaBoost

1. In this exercise we are given a decision stump for the following data:



To visualize it as a stump, suppose the decision threshold is T:



The current decision boundary fails on one occasion when classifying -1 as $+1$.

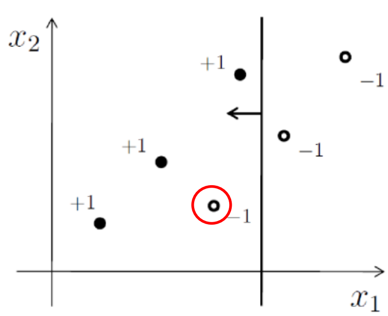
On the first iteration of AdaBoost each sample receives an equal weight, and since there are 6 samples each sample is weighted $\frac{1}{6}$. One sample was misclassified so the error is $\frac{1}{6}$.

$$\boldsymbol{\varepsilon_1 = \frac{1}{6}}$$

We use the following formula for finding the weight for the current stump:

$$\boldsymbol{a_1} = \frac{1}{2}\log\left(\frac{1-\varepsilon_1}{\varepsilon_1}\right) = \frac{1}{2}\log(5) \cong \mathbf{0.349}$$

2. Let's calculate the new weights for the samples, since one of the -1 samples got wrongly classified it would have a greater weight.



The sample that got misclassified.

We will use the formula to update the weights and then normalize it so the weights add up to 1.

$$d_{t+1}(x_i) = d_t(x_i) \cdot e^{-a_t y_i f_t(x_i)}$$

Basically, for every **correct** sample we update the weights like so:

$$d_2(x_i) = d_1(x_i) \cdot e^{-a_1} = \frac{1}{6} \cdot e^{-0.349} \cong 0.117$$

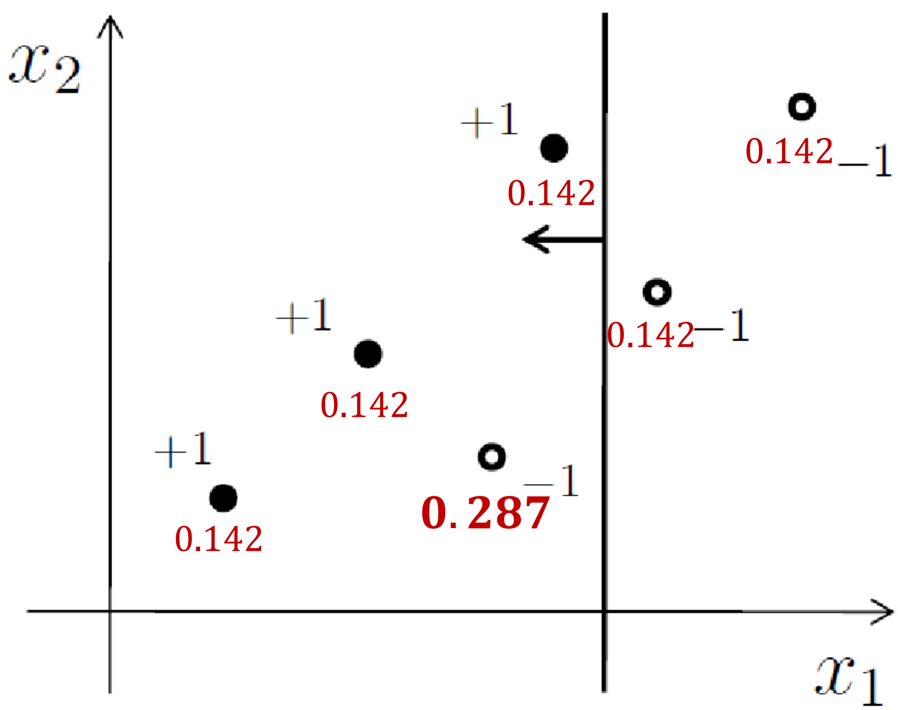
And for every **incorrect** sample we update the weights like so:

$$d_2(x_i) = d_1(x_i) \cdot e^{-a_1} = \frac{1}{6} \cdot e^{0.349} \cong 0.236$$

To normalize the new weights, we'll divide the total sum of weights out of each weight:

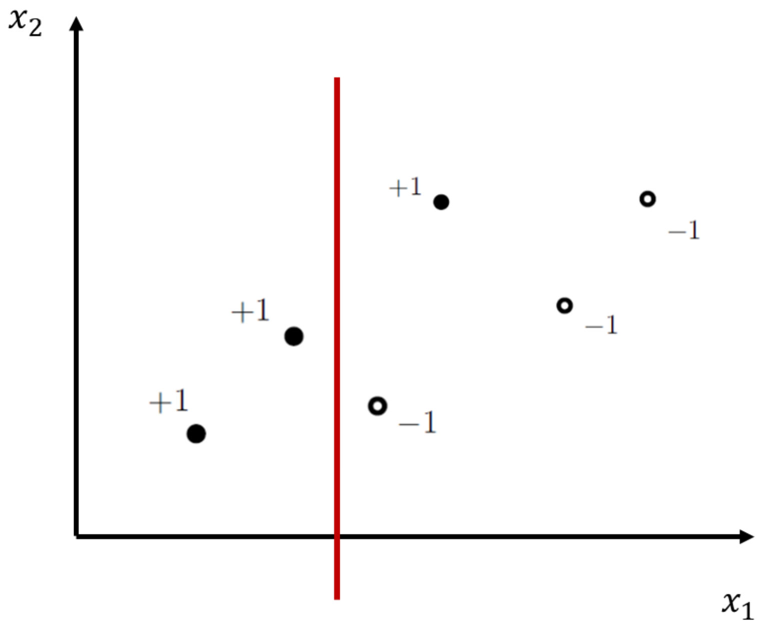
$$d_2(x) = 5 \cdot 0.117 + 1 \cdot 0.236 = 0.821$$

So, for every correctly classified sample the new weight is $\frac{0.117}{0.821} = 0.142$, and for the incorrectly classified sample the new weight is $\frac{0.236}{0.821} = 0.287$, depicted in the following figure:

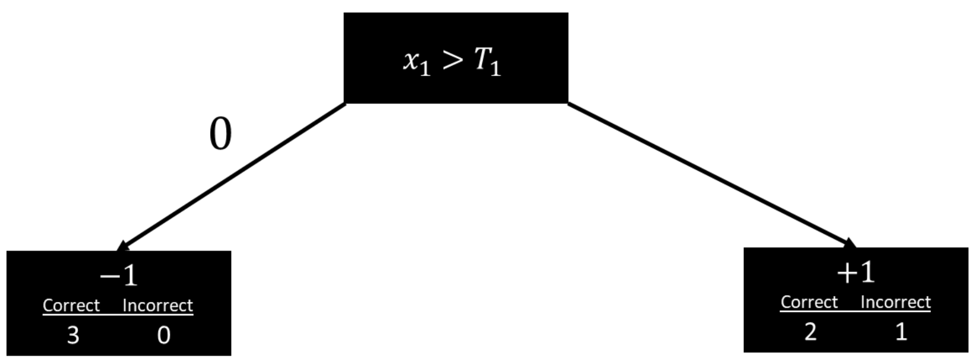


New weights for each sample added in red.

3. The new stump could use the following boundary:

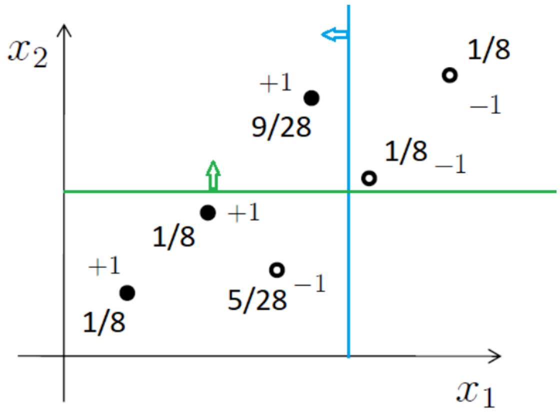


The stump would look like this:



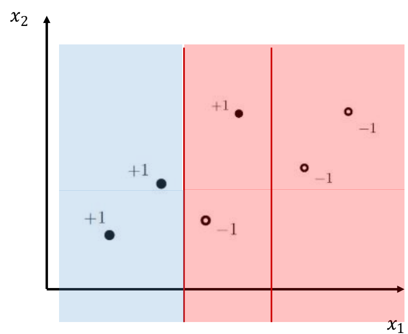
4. a_2 would have a larger classifier weight. ε_2 is going to be smaller than the first error since we incorrectly classified a sample with a lower weight than before, that means the formula for calculating a_2 would yield a greater result since it is a logarithmic formula.

5. It looks like the blue line would be a better decision stump because it has a lower error rate.



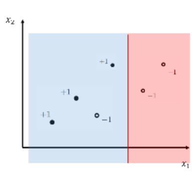
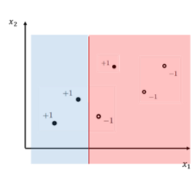
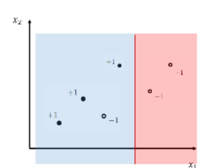
6. We know that: $a_1 = 0.349, a_2 = 1.1, a_3 = 0.62$, we can calculate the final classifier using:

$$a_1 * h_1 + a_2 * h_2 + a_3 * h_3 = 0.349 * h_1 + 1.1 * h_2 + 0.62 * h_3$$



In the figure the areas were decided by what classifier was 'heavier' in each of them.

For illustration purposes here is the calculation:

Sign($0.349 *$  $1.1 *$  $0.62 *$ )

The accuracy for the final classifier is $\frac{5}{6}$ since it misclassifies one sample.