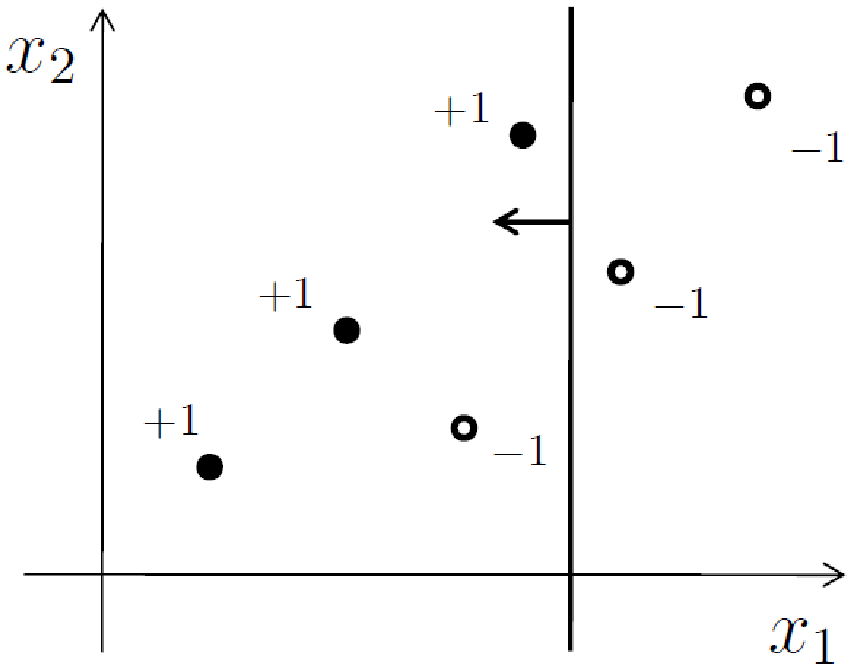
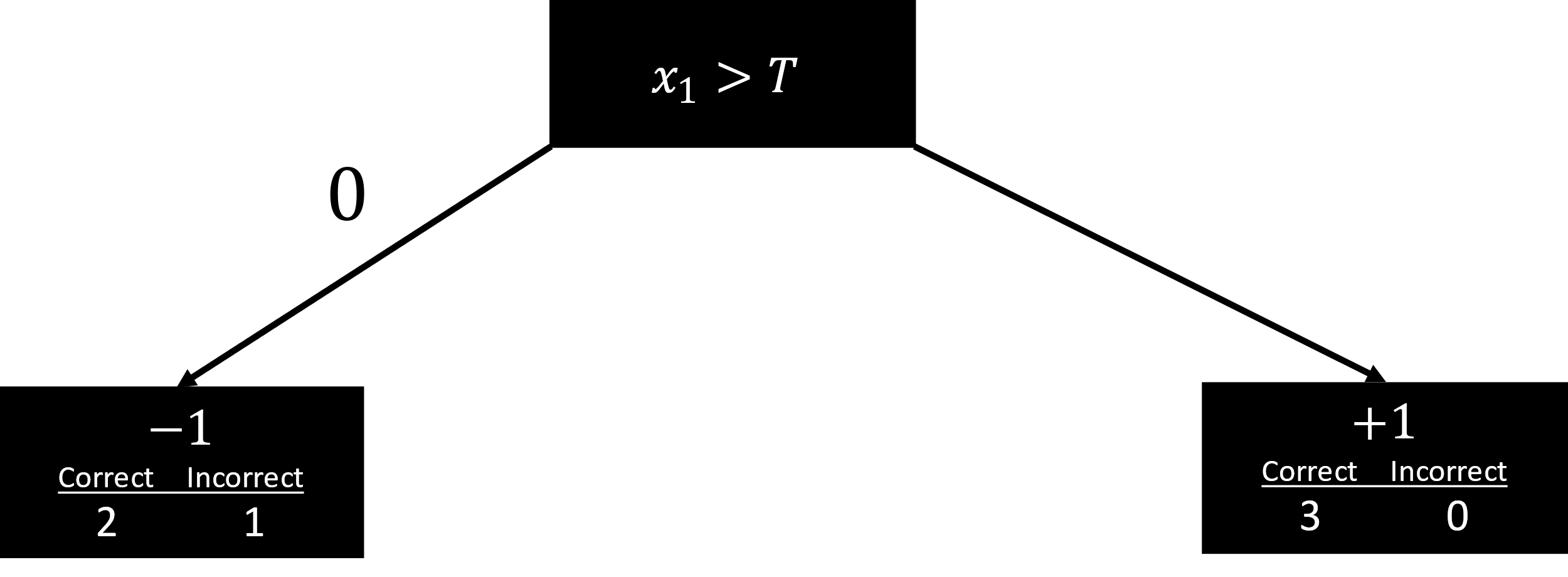
HW4 - Question 2 – AdaBoost

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 :

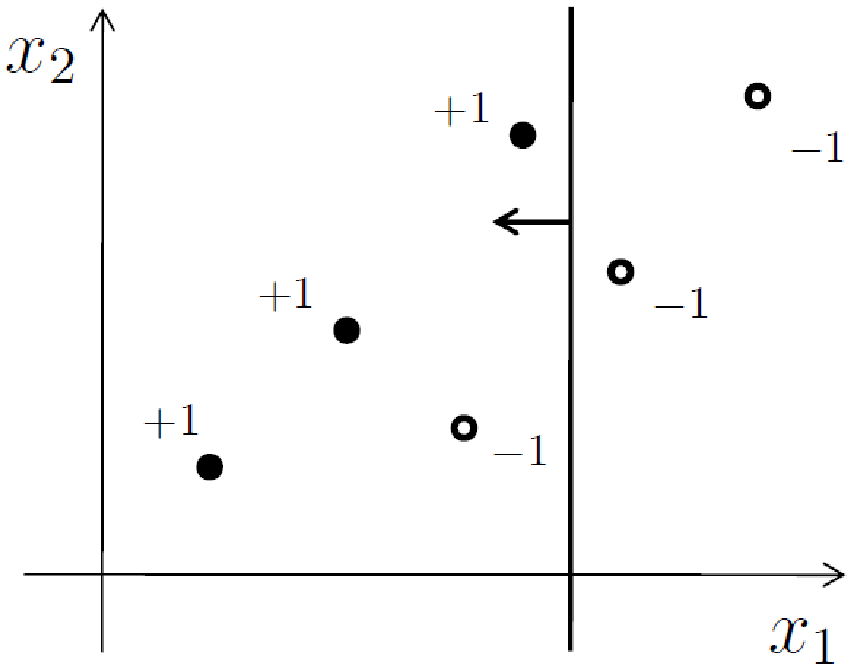


The current decision boundary fails on one occasion when classifying as .

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

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

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



The sample that got misclassified.

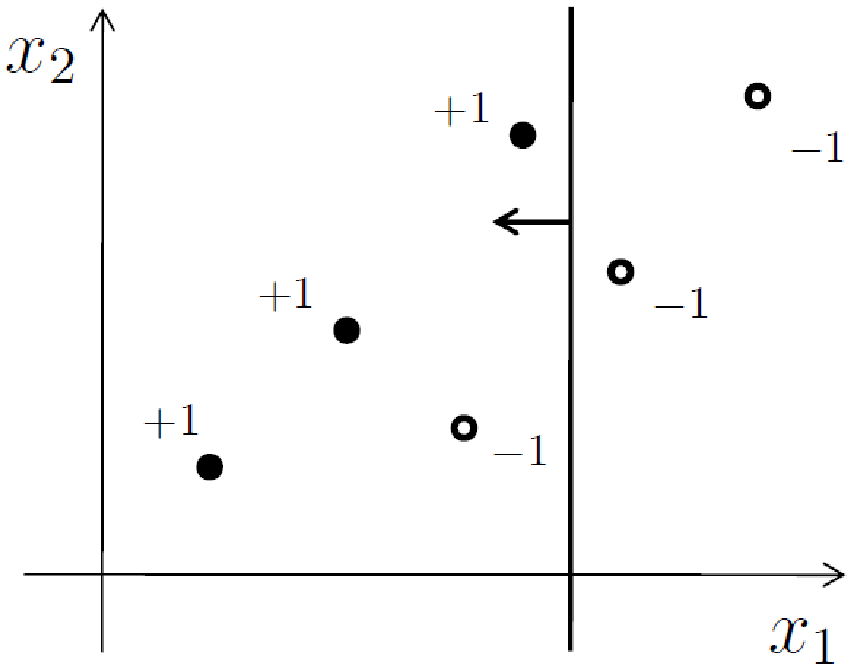
We will use the formula to update the weights and then normalize it so the weights add up to 1.

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

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

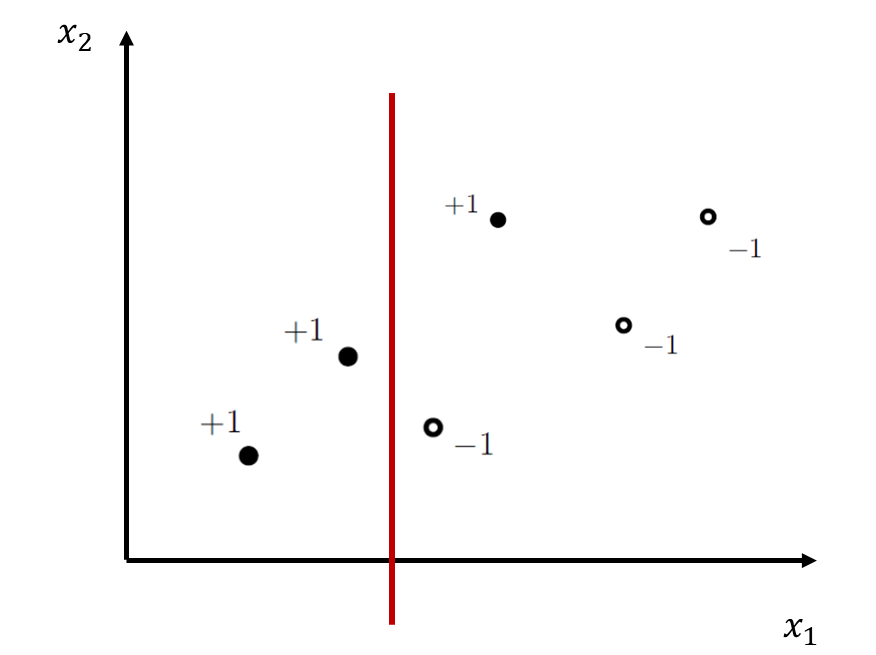
To normalize the new weights, we’ll divide the total sum of weights out of each weight:

So, for every correctly classified sample the new weight is , and for the incorrectly classified sample the new weight is , depicted in the following figure:

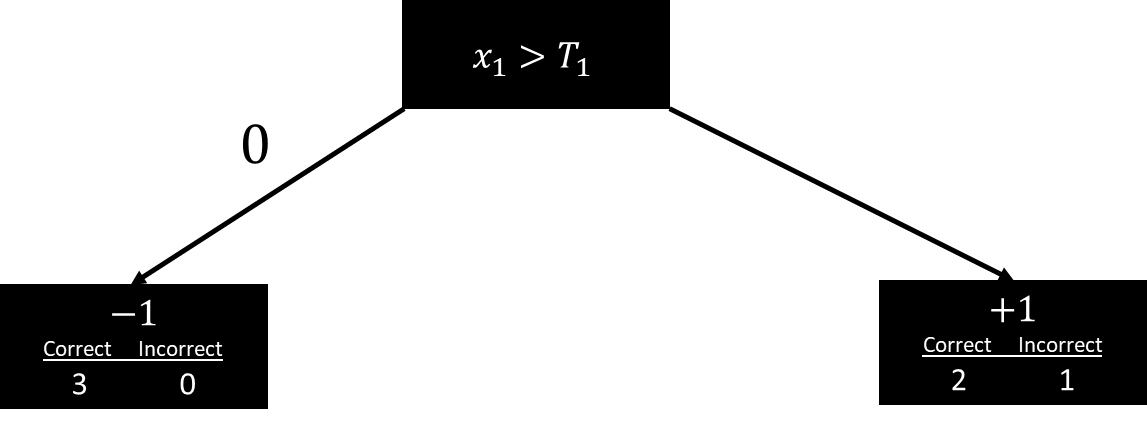


New weights for each sample added in red.

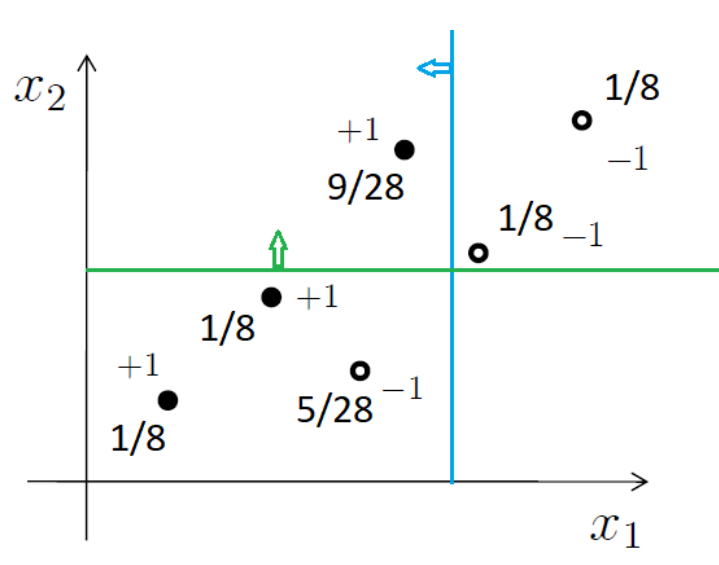
1. The new stump could use the following boundary:



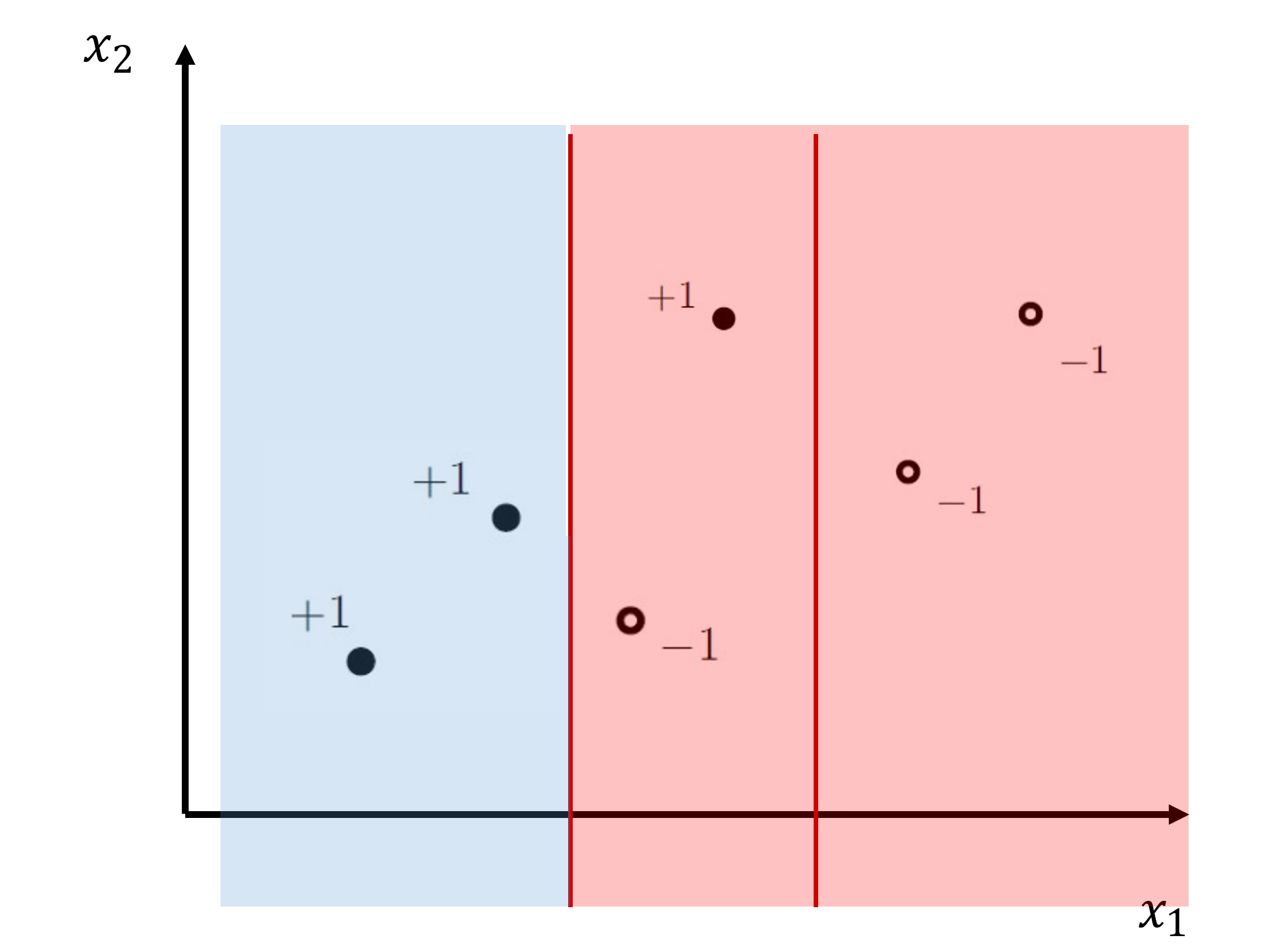
The stump would look like this:



1. would have a larger classifier weight. 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 would yield a greater result since it is a logarithmic formula.
2. It looks like the blue line would be a better decision stump because it has a lower error rate.



1. We know that: , we can calculate the final classifier using:



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

For illustration purposes here is the calculation:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Sign( |  |  |  |  |  |  | ) |

The accuracy for the final classifier is since it misclassifies one sample.