Boosting

1. Initial weight assigned to each data-point is = 1/m , Where m is the number of points
   $\implies$ weight $(D_1(i)) = 1/5 = 0.2$

2. Updated weights of each data point, based on the decision stump, are calculated as follows:

$$D_{t+1}(i) = \frac{D_t(i)exp(-\alpha_t y_i h_t(x_i))}{Z_t} \tag{1}$$

$$\alpha_t = \frac{1}{2}ln\Big(\frac{1-\epsilon_t}{\epsilon_t}\Big) \tag{2}$$

$$\epsilon_t = \sum_{i=1}^{m} D_t(i)\delta(h_t(x_i \neq y_i) \tag{3}$$

For the given decision stump:
$\epsilon_1 = 0.2 * 1 = 0.2$
$\alpha_1 = 0.5 * ln(0.8/0.2) = 0.693$

Updated weights for the correctly classified points are as follows:
$D_2(i) = D_1(i) * exp(-\alpha_1) = 0.2 * e^{-0.693} = 0.5/Z_t$

Updated weight for the mis classified point:
$D_2(i) = D_1(i) * exp(\alpha_1) = 0.2 * e^{0.693} = 2/Z_t$
Normalization, $Z_1 = 0.8$

Hence the updated weights are, **0.125** for the correctly classified points and **0.5** for the mis-classified point.

In the next iteration the decision stump is shifted as shown in the figure below:

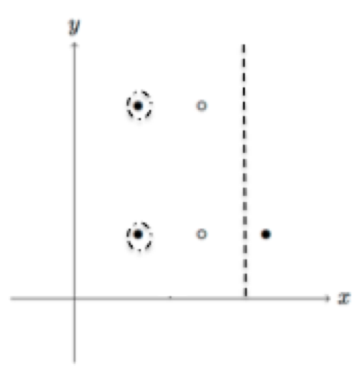

Figure 1

All the points to the right are classified as true(dark black) and the to the left are classified as false. Hence the points circled above are the ones which are mis-classified are the ones whose weights will increase and are circled.

3. No it is not possible to reach zero training error with increasing rounds of boosting. Based on the above figure we see that the error increases. Recalculating the weights would result in the mis-classified point with a weight of 0.25 the correctly classified points would get a weight of 0.083 and 0.33(for the point which was mis-classified in the first iteration). The error in this case would be higher. Thus with increasing number of iterations the error in this case would tend to 1.

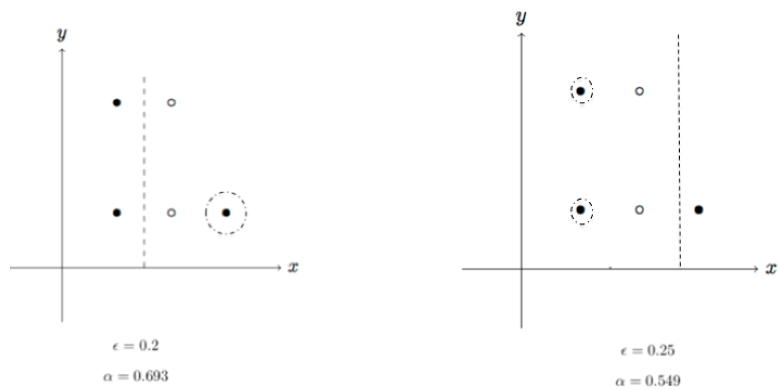

Figure 2: 1st and 2nd iterations