Adaboost Algorithm

Adaboost is a kind of boosting algorithm, the main idea of Adaboost is to combine a set of weak learners until certain condition, and such as low training error or certain number of weak classifiers (certain number of iteration times). The main algorithm includes three main steps:

1. Initialize the original weights. For all the N training samples, we only compare two of the different classes at one time, and assign {-1,+1} as labels to the two classes to show the different classes. The original values of the data's weight are all the same (1/n).
2. To train the weak classifiers. To be more specific, in the algorithm, we are more concern about the samples that are misclassified. To be more specific, after every iteration, if the samples have been classified correctly, their weights will be lower; otherwise, if a sample has be misclassified, its weight will be higher. The new weights are used for next iteration to help to choose.
3. Combine all the weak classifiers to be a strong classifier. After training, we add the weight of the classifier with smaller error rate, and make it has bigger weight in the final strong classifier.

The detailed procedures are like:
For a given training set D = {(x1,y1),(x2,y2),...(xn,yn)}, the xi represents the features, and yi is the labels which present two different classes, and since the domain of y is {-1,+1}

1. Initialize training set, and the original weight set $W_1(i) = 1/n$, n is the number of samples in the training set. And we choose a K_{max} as the iteration stop value;
2. for k = 1:K_{max}
3. train the weak classifier C_k using D sampled according to the weight W.
4. calculate the error E_k of C_k measured on D. The total error is the sum of their weight in W;
5. calculate $\alpha_k = \frac{1}{2}\log\frac{1-E_k}{E_k}$
6. update the weight W, by
   $W_{k+1}(i) = \frac{W_k(i)}{z_k} \times \begin{cases} e^{-\alpha_k} \\ e^{\alpha_k} \end{cases}$ if classified correct, multiply $e^{-\alpha_k}$,or otherwise
7. until k = K_{max}
8. return C_k and the correspondent α_k
9. end

The final classifier will be:

$g(x) = \left[\sum_{1}^{K_{\max}} \alpha_k h_k(x) \right]$, the $h_k(x)$ here also falls into the domain {-1,+1}, and indicates the class of present data.