AdaBoost Exercises

1. The AdaBoost algorithm can be summarized as follows:

During training, repeat the following for N iterations:

- Train a base C_j classifier on the weighted data set (initially, all weights are equal).
- Apply the classifier to all m training examples and calculate the weighted error (sum of the weights of the misclassified samples):

$$r_j = \sum_{\substack{i=1 \\ \hat{y}_j^{(i)} \neq y_j^{(i)}}}^{m} w_i$$

- Calculate the weight/importance of the new classifier:

$$\alpha_j = \ln \frac{1 - r_j}{r_j}$$

- Update the sample weights according to:

$$w_i^{(j+1)} = \begin{cases} w_i^{(j)} & \text{if } C_j(x_i) = y_i \\ e^{\alpha_j} w_i^{(j)} & \text{if } C_j(x_i) \neq y_i \end{cases}$$

- Normalize the weights so that they sum to 1:

$$w_i^{(j+1)} \leftarrow \frac{w_i^{(j+1)}}{Z_{j+1}}$$

where Z_j is the sum of the pre-normalized weights: $Z_{j+1} = \sum_{i=1}^{m} w_i^{(j+1)}$

At prediction, time select the class that gets the most (weighted) votes:

$$C^*(x) = \arg\max_k \sum_{\substack{j=1 \\ C_j(x)=k}}^{N} \alpha_j$$

Consider the following training set and initial weights:

x:	.1	.3	.6	.9
y:	-1	1	1	-1
$\mathbf{w}_0$:	.25	.25	.25	.25

(a) Assume that C_0 has been created, with the following result:

x:	.1	.3	.6	.9
$\hat{y}_0$:	1	1	1	-1
correct?	N	Y	Y	Y

What is r_0? What is α_0? Show your work.

(b) Fill in the table below with the updated weights. Show your work.

x:	.1	.3	.6	.9
$\mathbf{w}_1$:				

(c) Now assume that C_1 performs as follows:

x:	.1	.3	.6	.9
$\hat{y}_1$:	-1	1	1	1
correct?	Y	Y	Y	N

Calculate r_1, α_1 and the updated weights:

x:	.1	.3	.6	.9
$\mathbf{w}_2$:				

(d) Given that C_0 classifies a particular point as -1 and C_1 classifies it as 1, what class will be selected by the ensemble? Justify your answer.