Consider a dataset given in the following:

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
y	1	1	1	-1	-1	-1	-1	1	1	1

Similar to Bagging, we can pick different training examples to obtain a new training set in AdaBoost. Supposed we have three rounds of boosting and each round has the following training record respectively:

Round 1

x	0.1	0.4	0.5	0.6	0.6	0.7	0.7	0.7	0.8	1
y	1	-1	-1	-1	-1	-1	-1	-1	1	1

Round 2

x	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3
y	1	1	1	1	1	1	1	1	1	1

Round 3

x	0.2	0.2	0.4	0.4	0.4	0.4	0.5	0.6	0.6	0.7
y	1	1	-1	-1	-1	-1	-1	-1	-1	-1

The best split point for each round is:

Round	**Split Point**	**Left Class**	**Right Class**
1	0.75	-1	1
2	0.05	1	1
3	0.3	1	-1

Use AdaBoost and follow what we have covered in the toy example to compute ε_i, α_i and the updated weights. Finally find the combined classifier H.

從 3 個 boosting round 以及題目所指定的 best split point 可以得到下表的結果，與題目所給的 dataset 相比，可以看出每個 round 都有被分錯的地方，以紅色去標記：

	Original Data	**x**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
		y	**1**	**1**	**1**	**-1**	**-1**	**-1**	**-1**	**1**	**1**	**1**
Round	1		-1	-1	-1	-1	-1	-1	-1	1	1	1
	2		1	1	1	1	1	1	1	1	1	1
	3		1	1	1	-1	-1	-1	-1	-1	-1	-1

$$a_j = \frac{1}{2}\ln\left(\frac{1-\varepsilon_j}{\varepsilon_j}\right)$$

Round 1

On round 1, AdaBoost assigns equal weight to all of the examples. Given examples with these weights, the base learner chooses the base hypothesis indicated by h_1, this hypothesis incorrectly classifies three points so its error ε_1 is 0.30、$\alpha_1 \approx 0.42$、$Z_1 \approx 0.92$。

	1	2	3	4	5	6	7	8	9	10
$D_1(i)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$e^{-\alpha_1 y_i h_1(x_i)}$	1.53	1.53	1.53	0.66	0.66	0.66	0.66	0.66	0.66	0.66
$D_1(i)e^{-\alpha_1 y_i h_1(x_i)}$	0.15	0.15	0.15	0.07	0.07	0.07	0.07	0.07	0.07	0.07
$D_2(i)$	0.17	0.17	0.17	0.07	0.07	0.07	0.07	0.07	0.07	0.07

Round 2

On round 2, the base learner pays special attention on the three relatively high-weight points missed by h_1 and chooses h_2, under distribution D_2, these four points have weight only around 0.07143, so the error of h_2 with respect to D_2, $\varepsilon_2 \approx 0.29$、$\alpha_2 \approx 0.46$ and $Z_2 \approx 0.90$.

	1	2	3	4	5	6	7	8	9	10
$D_1(i)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
$e^{-\alpha_1 y_i h_1(x_i)}$	1.53	1.53	1.53	0.66	0.66	0.66	0.66	0.66	0.66	0.66
$D_1(i)e^{-\alpha_1 y_i h_1(x_i)}$	0.15	0.15	0.15	0.07	0.07	0.07	0.07	0.07	0.07	0.07
$D_2(i)$	0.17	0.17	0.17	0.07	0.07	0.07	0.07	0.07	0.07	0.07
$e^{-\alpha_2 y_i h_2(x_i)}$	0.63	0.63	0.63	1.58	1.58	1.58	1.58	0.63	0.63	0.63
$D_2(i)e^{-\alpha_2 y_i h_2(x_i)}$	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.05	0.05	0.05
$D_3(i)$	0.12	0.12	0.12	0.16	0.16	0.16	0.16	0.05	0.05	0.05

Round 3

On round 3, this classifier misses none of the points misclassified by h_1 and h_2 since these points have relatively high weight under D_3. Instead, it misclassifies three points, they were not misclassified by h_1 or h_2, so they have very low weight under D_3. On round 3, $\varepsilon_3 \approx 0.15$、$\alpha_3 \approx 0.87$ and $Z_3 \approx 0.71$.

	1	2	3	4	5	6	7	8	9	10	
$D_1(i)$	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	$\varepsilon_1 \approx 0.30$
$e^{-\alpha_1 y_i h_1(x_i)}$	1.53	1.53	1.53	0.66	0.66	0.66	0.66	0.66	0.66	0.66	$\alpha_1 \approx 0.42$
$D_1(i)e^{-\alpha_1 y_i h_1(x_i)}$	0.15	0.15	0.15	0.07	0.07	0.07	0.07	0.07	0.07	0.07	$Z_1 \approx 0.92$
$D_2(i)$	0.17	0.17	0.17	0.07	0.07	0.07	0.07	0.07	0.07	0.07	$\varepsilon_2 \approx 0.29$
$e^{-\alpha_2 y_i h_2(x_i)}$	0.63	0.63	0.63	1.58	1.58	1.58	1.58	0.63	0.63	0.63	$\alpha_2 \approx 0.46$
$D_2(i)e^{-\alpha_2 y_i h_2(x_i)}$	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.05	0.05	0.05	$Z_2 \approx 0.90$
$D_3(i)$	0.12	0.12	0.12	0.16	0.16	0.16	0.16	0.05	0.05	0.05	$\varepsilon_3 \approx 0.15$
$e^{-\alpha_3 y_i h_3(x_i)}$	0.42	0.42	0.42	0.42	0.42	0.42	0.42	2.38	2.38	2.38	$\alpha_3 \approx 0.87$
$D_3(i)e^{-\alpha_3 y_i h_3(x_i)}$	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.12	0.12	0.12	$Z_3 \approx 0.71$

Compute ε_i, α_i

	ε_i	α_i
Round1	0.30	0.42
Round2	0.29	0.46
Round3	0.15	0.87

Combined classifier *H*

Original Data	**x**	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1**
	y	**1**	**1**	**1**	**-1**	**-1**	**-1**	**-1**	**1**	**1**	**1**
1		-1	-1	-1	-1	-1	-1	-1	1	1	1
2		1	1	1	1	1	1	1	1	1	1
3		1	1	1	-1	-1	-1	-1	-1	-1	-1
$\alpha_i y_{round_i}$		0.90	0.90	0.90	-0.83	-0.83	-0.83	-0.83	0.01	0.01	0.01

$$\alpha_1 y_{round_1} + \alpha_2 y_{round_2} + \alpha_3 y_{round_3} > 0 \text{ , } y_n = 1$$
$$\alpha_1 y_{round_1} + \alpha_2 y_{round_2} + \alpha_3 y_{round_3} < 0 \text{ , } y_n = -1$$

Combined classifier *H*：

$$H = \text{sign}(\alpha_1 y_{round_1} + \alpha_2 y_{round_2} + \alpha_3 y_{round_3})$$

藉由3個round所得出的權重α_1、α_2、α_3，可以得到最後每個x所對應的值，若所算出的數字大於0，則y就會是1相反地，若數字小於0，y就為-1。與原始資料相比二者相同。

$\sum_{n=1}^{3} \alpha_i y_n$	0.90	0.90	0.90	-0.83	-0.83	-0.83	-0.83	0.01	0.01	0.01
y(Adaboost)	1	1	1	-1	-1	-1	-1	1	1	1