Algorithm AdaBoost.M1

Input:

- Sequence of N examples S = [(x_i , y_i)], $i = 1, \dots, N$ with labels $y_i \in \Omega$, $\Omega = \{ w_1 , \dots , w_C \}$;
- Weak learning algorithm **WeakLearn**;
- Integer T specifying number of iterations

Initialize $D_1(i) = \frac{1}{N}$., $i = 1, \dots, N$

Do for $t = 1, 2, \dots, T$:

1. Select a training data subset S_t , drawn from the distribution D_t.
2. Train **WeakLearn** with S_t , receive hypothesis h_t.
3. Calculate the error of

   $$h_t: \varepsilon_t = \sum_{t h_t(x_i) \neq y_i} D_t(i).$$

   If $\varepsilon > \frac{1}{2}$, **abort**.
4. Set $\beta_t = \varepsilon_t/(1 - \varepsilon_t)$.
5. Update distribution

   $$D_t: D_{t+1}(i) = \frac{D_t(i)}{Z_t} \; x \begin{cases} \beta_t & if \; h_t(x_i) = y_i \\ 1, & otherwise \end{cases}$$

   where $Z_t = \sum_i D_t(i)$ is a normalization constant chosen so that D_{t+1} becomes a proper distribution function.

Test – Weighted Majority Voting: Given an unlabeled instance x,

1. Obtain total vote received by each class

   $$V_j = \sum_{t: h_i(x) = \omega_j} \log\left(\frac{1}{\beta_t}\right), j = 1, \dots, C.$$

2. Choose the class that receives the highest total vote as the final classification.

i =>	**0_1**	**0_2**	**0_3**	**0_4**	**0_5**	**0_6**	**0_7**	**0_8**	**0_9**	**0_{10}**
D_1	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$
h_1			x	x	x				x	
misclassified ($h1(0_i) \neq y_i$)										
$\varepsilon_1 =$	$\frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = \frac{4}{10}$					$\beta_1 = \frac{\varepsilon_1}{(1-\varepsilon_1)} = \frac{\left(\frac{4}{10}\right)}{\left(\frac{6}{10}\right)} = \frac{2}{3}$ $\frac{1}{\beta_1} = \frac{3}{2}$ $(less\ confidence)$				
D_2	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{1}{10}$	$\frac{1}{15}$
Normalize	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{12}$
		0_2	0_3	0_4		0_6	0_7		0_9	
h_2				x						
$\varepsilon_2 =$	$\frac{1}{8}$					$\beta_2 = \frac{\varepsilon_2}{1-\varepsilon_2} = \frac{\frac{1}{8}}{\frac{7}{8}} = \frac{1}{7}$ $\frac{1}{\beta_2} = 7$ $(more\ confidence)$				
D_3	$\frac{1}{12}$	$\frac{1}{12} * \frac{1}{7}$	$\frac{1}{8} * \frac{1}{7}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{12} * \frac{1}{7}$	$\frac{1}{12} * \frac{1}{7}$	$\frac{1}{12}$	$\frac{1}{8} * \frac{1}{7}$	$\frac{1}{12}$
	$\frac{1}{12}$	$\frac{1}{84}$	$\frac{1}{56}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{84}$	$\frac{1}{84}$	$\frac{1}{12}$	$\frac{1}{56}$	$\frac{1}{12}$
Normalize D_3:	$\frac{588}{7056}$	$\frac{84}{7056}$	$\frac{126}{7056}$	$\frac{441}{7056}$	$\frac{441}{7056}$	$\frac{84}{7056}$	$\frac{84}{7056}$	$\frac{588}{7056}$	$\frac{126}{7056}$	$\frac{588}{7056}$

For D2:

normalization: $z_1=\frac{\frac{1}{15}}{(\sum D_1(i))}=\frac{\frac{1}{15}}{\left(\frac{1}{15}+\frac{1}{15}+\frac{1}{10}+\frac{1}{10}+\frac{1}{10}+\frac{1}{15}+\frac{1}{15}+\frac{1}{15}+\frac{1}{10}+\frac{1}{15}\right)}=\frac{\frac{1}{15}}{\frac{4}{5}}=\frac{1}{12}$

$z_2=\frac{\frac{1}{10}}{(\sum D_1(i))}=\frac{\frac{1}{10}}{\left(\frac{1}{15}+\frac{1}{15}+\frac{1}{10}+\frac{1}{10}+\frac{1}{10}+\frac{1}{15}+\frac{1}{15}+\frac{1}{15}+\frac{1}{10}+\frac{1}{15}\right)}=\frac{\frac{1}{10}}{\frac{4}{5}}=\frac{1}{8}$

For D3:

$\frac{\frac{1}{12}}{(\sum D_1(i))}=\frac{\frac{1}{12}}{\frac{1}{12}+\frac{1}{84}+\frac{1}{56}+\frac{1}{8}+\frac{1}{8}+\frac{1}{84}+\frac{1}{84}+\frac{1}{12}+\frac{1}{56}+\frac{1}{12}}=\frac{588}{7056}$, $\frac{\frac{1}{84}}{(\sum D_1(i))}=\frac{84}{7056}$,

$\frac{\frac{1}{56}}{(\sum D_1(i))}=\frac{126}{7056}$, $\frac{\frac{1}{8}}{(\sum D_1(i))}=\frac{441}{7056}$