LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Database Systems Group • Prof. Dr. Thomas Seidl

Exercise 11: Adaboost

Knowledge Discovery in Databases I

SS 2016

AdaBoost - Algorithm

Initialize $w_1, \dots, w_N = \frac{1}{N}$ $(N: number\ of\ training\ instances)$

For $m = 1, \dots, M$:

- Fit a classifier $K_m(x)$ to the training data by minimizing weighted error function $J_m = \sum_{n=1}^{N} w_n I(K_m(x_n) \neq t_n)$ (t_n: correct class label, I: indicator function)
- Compute weighting coefficient $\alpha_m = \ln\left\{\frac{1-\epsilon_m}{\epsilon_m}\right\}$ with $\epsilon_m = \frac{J_m}{\sum_1^N w_n}$
- Updata all data weights: $w_{n,old} = w_n$

  $w_n = w_{n,old} \exp\{\alpha_m I(K_m(x_n) \neq t_n)\}$

Make prediction using final model

$$K_M(x) = sign\left(\sum_{m=1}^{M} \alpha_m K_m(x)\right)$$

Exercise 11-1: Adaboost

$\epsilon_m = \frac{J_m}{\sum_1^N w_n}$	$\frac{1-\epsilon_m}{\epsilon_m}$	$\alpha_m = \ln\left\{\frac{1-\epsilon_m}{\epsilon_m}\right\}$	$w_n = w_{n,old} \exp\{\alpha_m I(K_m(x_n) \neq t_n)\}$			
			$w_{n,old} = 0.01$		$w_{n,old} = 0.1$	
			$I(K_m(x_n) \neq t_n) = 1$ (wrongly classified)	$I(K_m(x_n) \neq t_n) = -1$ (correctly classified)	$I(K_m(x_n) \neq t_n) = 1$ (wrongly classified)	$I(K_m(x_n) \neq t_n) = -1$ (correctly classified)
0.01	99	4.6	$e^{4.6} = 99.5$ $w_n = 99.5 * 0.01$ $= 0.995$	$e^{-4.6} = 0.01$ $w_n = 0.01 * 0.01$ $= 0.0001$	$e^{4.6} = 99.5$ $w_n = 99.5 * 0.1$ $= 9.95$	$e^{-4.6} = 0.01$ $w_n = 0.01 * 0.1$ $= 0.001$
0.5	1	0	$e^{0} = 1$ $w_n = 1 * 0.01$ $= 0.01$	$e^{-0} = 1$ $w_n = 1 * 0.01$ $= 0.01$	$e^{0} = 1$ $w_n = 1 * 0.1$ $= 0.01$	$e^{-0} = 1$ $w_n = 1 * 0.1$ $= 0.01$
0.99	0.01	-4.6	$e^{-4.6} = 0.01$ $w_n = 0.01 * 0.01$ $= 0.0001$	$e^{4.6} = 99.5$ $w_n = 99.5 * 0.01$ $= 0.995$	$e^{-4.6} = 0.01$ $w_n = 0.01 * 0.1$ $= 0.001$	$e^{4.6} = 99.5$ $w_n = 99.5 * 0.1$ $= 9.95$

$$K_M(x) = sign\left(\sum_{m=1}^{M} \alpha_m K_m(x)\right)$$