Kiran Bhat Gopalakrishna
kxb140230

1) Boosting

1000 total training examples

Alg A → classifier A → 800 correct, 200 incorrect

Alg B → classifier B → 800 correct, 200 incorrect.
All incorrect of B are correct in A

Alg C → classifier C → 900 correct, 100 incorrect
All incorrect of C are correct on A & B

a) Weight of classifier A

$D_1(i) = 1/1000$

$\epsilon_1 = 200 * D_1(i)$

$= 200 \times \frac{1}{1000} = 0.2$

Choose $\alpha_t = \frac{1}{2} \ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$

$\alpha_1 = \frac{1}{2} \ln\left(\frac{1-\epsilon_1}{\epsilon_1}\right) = \frac{1}{2} \ln\left(\frac{1-0.2}{0.2}\right) = 0.69314$

After the update, instance correctly classified as

$$D_2(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-\alpha_t} & \text{if } h_t(x_i) = y_i \\ e^{\alpha_t} & \text{if } h_t(x_i) \neq y_i \end{cases}$$

$$= \frac{D_t(i) \exp(-\alpha_t y_i h_t(x_i))}{Z_t}$$

Here Z_t is a normalization factor (so that D_{t+1} is distribution

$$D_2(i) = \frac{D_1(i)}{Z_t} e^{-\alpha t} = \frac{0.001}{0.8} \times e^{-0.69314}$$

$$= 6.25\, e^{-t}$$

others as

$$D_2(i) = \frac{D_1(i)}{Z_t} e^{\alpha t} = \frac{0.001}{0.8} \times e^{0.69314}$$

$$= 2.499 \times 10^{-03}$$

Weight of classifier is $\alpha_t = \frac{1}{2} \ln \frac{(1-\epsilon_1)}{\epsilon_1}$

$$= 0.69314$$

b) Classification result for $A=1, B=0$ error $=200$

$A=0, B=1$ error $=200$

$A=1, B=1$ error $=600$

$$e_2 = 6.25\, e^{-9} \times 200 = 0.125$$

$$\alpha_2 = \frac{1}{2} \ln \frac{(1-\epsilon_1)}{\epsilon_1} = 0.9729$$

$$Z_t = 0.6914$$

Instances with error 2

$$D(i) = \frac{D_2(i)}{Z_t} e^{-\alpha t} = \frac{0.00094491}{Z_t} = 0.001t$$

Instances with error 3

$$D(i) = \frac{D_2(i)}{Z_t} e^{-\alpha t} = \frac{0.00023}{Z_t} = 0.00357$$

Weight for classifier

$$\alpha_2 = \frac{1}{2} \ln \frac{(1-\epsilon_1)}{\epsilon_1} = 0.9729$$

c) $A=0 \quad B=1 \quad C=1 \quad error_1 = 200$

$A=1 \quad B=0 \quad C=1 \quad error_2 = 200$

$A=1 \quad B=1 \quad C=1 \quad error_3 = 500$

$A=1 \quad B=1 \quad C=0 \quad error_4 = 100$

$$\epsilon = 0.0357$$

$$\alpha_3 = \frac{1}{2} \ln \frac{(1-\epsilon)}{\epsilon} = 1.64$$

= weight of classifier C

d) $A=1 \quad B=1 \quad C=-1$

$$h = \alpha_1 * 1 + \alpha_2 * 1 + \alpha_3 * -1$$

$$= 0.0180$$

∴ Positive

e) $A=1, \quad B=-1, \quad C=1$

$$h = \alpha_1 * 1 + \alpha_2 * -1 + \alpha_3 * 1$$

$$= 1.367$$

∴ Positive

f) $A = -1$, $B = 1$, $C = 1$

$h = \alpha_1 * -1 + \alpha_2 * 1 + \alpha_3 * 1$

$= 1.972$

$\therefore$ Positive

2) PAC learning

Integer set $X = \{1, 2, 3, 4 \ldots 50\}$

each hypothesis $h \in H$ is an interval of the form

$$a \leq x \leq b$$

where a & b are integers between 1 and 50 also $a \leq b$

Hypothesis labels instance X positive if it falls into the interval and negative if outside.

i) Distinct hypothesis in H

when $b = 50$ there are exactly 50 possible values for a

$\therefore \frac{(1+50) \times 50}{2} = \frac{51 \times 50}{2} = 1275$

ii) Probability = 0.95

True error = 0.1

Assuming instances are drawn at random and as per probability distribution $P(x)$ then each training example would be generated by drawing an instance at random according to $P(x)$ and labelling it

$$m \geq \frac{1}{\epsilon}\left(\ln|H| + \ln\frac{1}{\delta}\right)$$

where H - distinct hypothesis

ϵ - error Rate

$(1-\delta)$ - Probability

$$= \frac{1}{0.1}\left(\ln|1275| + \ln\frac{1}{1-0.95}\right)$$

$$= \frac{1}{0.1}(7.150 + 2.99)$$

$$= 101.457$$

3) VC Dimensions

VC dimension = 2

A) If a target function is specified by an interval and labelling is done as positive iff it lies inside the interval then set of 2 points can be shattered since there is only single block of positive examples that could lie within the interval. (++) - can be shattered

However if there are 3 set of points, it can not be shattered since it can not be labelled in +, −, + order

B) Axis Aligned Rectangle in the plane R^2

VC-dimension = 4

For instances, consider points that form rectangle.
say $n=5$ points $\{v_1, v_2, v_3, v_4 v_5\} \in R^2$

Rectangle contains the points with max x-coordina
min x wordinate, max-y & min-y coordinates respectivel
∴ atmost 4 points

let $S \subset \{v_1, v_2, v_3, v_4 v_5\}$

∴ Any axis-aligned rectangle that contains S must also contain all of the points. v_1 v_2 v_3 v_4 and v_5 and there is one V_p that is not in S, but still must be in a rectangle

Hence even if the smallest enclosing box is drawn around the 5 points it is not possibl to label the point inside the box and the remaining points on the edges.