Name SOLUTIONS

Quiz 3

Please write all answers on these pages.

Useful formulas for Adaboost:

$$\varepsilon_t = \sum_{j=1}^{N} \mathbf{w}_t(j)\, \delta(y_j \neq h_t(\mathbf{x}_j)), \text{ where}$$

$$\delta(y_j \neq h_t(\mathbf{x}_j)) = \begin{cases} 1 \text{ if } y_j \neq h_t(\mathbf{x}_j) \\ 0 \text{ otherwise} \end{cases}$$

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

1. Suppose you have trained three support vector machines, h_1, h_2, and h_3, returning binary classifications (+1 or −1). The observed accuracy of each of the hypotheses is 55%. Assuming that the errors of these hypotheses are independent, what is the predicted accuracy of an ensemble hypothesis H, where H's classification of an instance is a majority vote of the classifications of h_1, h_2, and h_3? **(2 points)**

H = majority vote of h_1, h_2, h_3

Cases where H is Correct

h_1	h_2	h_3	Probability
Correct	Correct	Correct	$(.55)^3 = .166$
Correct	Correct	Incorrect	$(.55)^2 \cdot (.45) = .136$
Correct	Incorrect	Correct	.136
Incorrect	Correct	correct	.136
		Total	

Probability that H is correct $\approx .57$

2. Suppose you have done one iteration of Adaboost to produce classifier h_1 and find that h_1 has the following results on the training data. (Assume the initial weights on the training data are uniform.) **(5 points)**

Example	True Class	Predicted Class ($h_1(\mathbf{x})$)
$\mathbf{x}_1$	+1	+1
$\mathbf{x}_2$	+1	−1
$\mathbf{x}_3$	−1	+1

(a) What is ε_1 ?

$w_1 = (1/3, 1/3, 1/3)$

$\varepsilon_1 = \frac{2}{3}$

(b) What is α_1 ?

$\alpha_1 = \frac{1}{2}\ln\left(\frac{1/3}{2/3}\right) = \frac{1}{2}\ln\left(\frac{1}{2}\right) = -.35$

(c) Now suppose that a second iteration of Adaboost was done and produced classifier h_2, which has the following results on the training data:

Example	True Class	Predicted Class ($h_1(\mathbf{x})$)
$\mathbf{x}_1$	+1	+1
$\mathbf{x}_2$	+1	+1
$\mathbf{x}_3$	−1	+1

$\hat{w}_2(1) = \frac{1}{3}\exp(+.35) = .47$

$\hat{w}_2(2) = \frac{1}{3}\exp(-.35) = .23$

$\hat{w}_3(3) = .23$

$Z = .47 + .23 + .23 = .93$

$w_2 = (.50, .25, .25)$

What is ε_2 ?

$\varepsilon_2 = .25$

(d) What is α_2 ?

$\alpha_2 = \frac{1}{2}\ln\left(\frac{3/4}{1/4}\right) = \frac{1}{2}\ln(3) = .55$

(e) At this point you stop iterating Adaboost, and, using h_1 and h_2, construct an ensemble classifier H. I give you a test example $\mathbf{x}_4$, and you find that $h_1(\mathbf{x}_4) = +1$ and $h_2(\mathbf{x}_4) = -1$. What is $H(\mathbf{x}_4)$? Show how you calculated the classification.

$H(x_4) = \text{sgn}(-.35(1) + .55(-1)) = -1$

3. Suppose you train an SVM using a linear kernel and find it has accuracy of .6 on the test set. You suspect the main source of the error is bias. Which of the following might help solve the bias problem? (Choose one.) **(1 point)**

(a) Increasing the size of the training set

(b) Using a more complex kernel (e.g., polynomial)

(c) Using a different test set

4. Now, suppose you find that error is not due to bias, so you suspect it is due to variance. Which of the following might help solve the variance problem? (Choose one.) **(1 point)**

(a) Increasing the size of the training set

(b) Using a more complex kernel (e.g., polynomial)

(c) Using a different test set