COMS 4771 Machine Learning (Spring 2015)
Problem Set #1

Nandini Shah - nss2158@columbia.edu
Discussants: kvm2116,kk3004

February 8, 2015

Problem 1

1.
Train error % = 14
Test error % = 15.8

2.
Train error % = 90.1283
Test error % = 90.2000
The OCR covariance matrix is not invertible. There are a lot of zeros and redundant dimensions. Matlab tries to invert it by calculating the 1/0 = NaN and produces warnings.The result is no predictions; all the labels are originally zeros and hence on comparison to the actual labels we appear to get approximately 90% error.

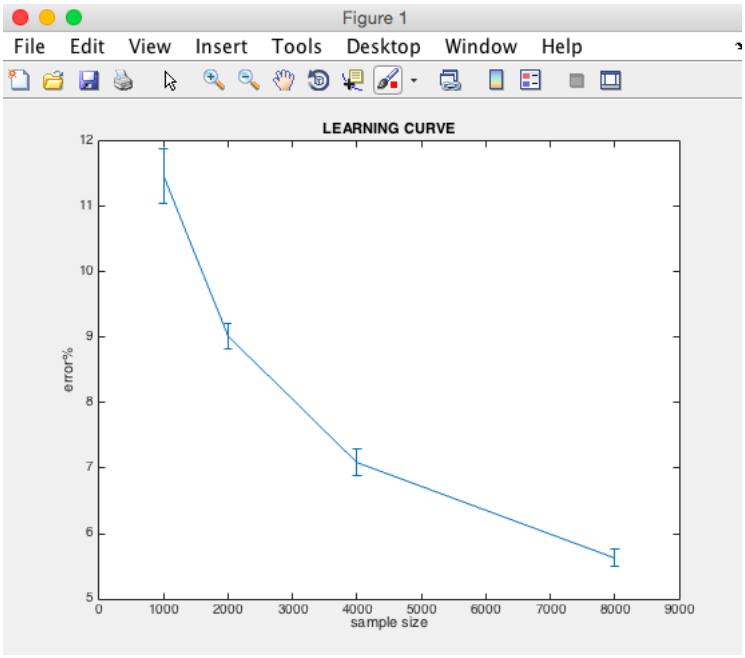
3.
Test error % = 17.9700
Train error % = 19.2017

Problem 2

For samples of size 1000, 2000, 4000, 8000 respectively,

Mean errors % = 11.47%, 9.02%, 7.09%, 5.63%

Std deviation = 0.0041, 0.0019, 0.0020, 0.0013



Problem 3

P$(Y = 0) = \frac{2}{3}$, P$(Y = 1) = \frac{1}{3}$
Class 0 has distribution N$(0, 1)$, Class 1 has distribution N$(1, \frac{1}{4})$
Penalty for 'false positive' = \$c
Penalty for 'false negative' = \$1

Current classifier:

$$f^*(x) = \begin{cases} 0 & \text{if } x \leq \text{b} \\ 1 & \text{if } x > \text{b} \end{cases}$$

where 'b' is the current threshold.

$$E(penalty) = c.P(y = 0|x > b) + 1.P(y = 1|x < b)$$

$$E(penalty) = c[P(y = 0).P(x > b|y = 0)] + 1.[P(y = 1).P(x < b|y = 1)] \rightarrow eqn.1$$

Now,

$$P(x > b|y = 0) = \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{\frac{-(x-\mu)^2}{2\sigma_y^2}}$$

$$P(x > b|y = 0) = \int_b^\infty \frac{1}{\sqrt{2\pi.1^2}} e^{\frac{-(x-0)^2}{2.1^2}} \mathrm{d}x$$

$$P(x > b|y = 0) = \int_b^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-(x)^2}{2}} \mathrm{d}x \rightarrow eqn.2$$

and

$$P(x \leq b|y = 1) = \int_{-\infty}^b \frac{1}{\sqrt{2\pi.\frac{1}{4}}} e^{\frac{-(x-1)^2}{2.\frac{1}{4}}} \mathrm{d}x$$

$$P(x \leq b|y = 1) = \int_{-\infty}^b \frac{2}{\sqrt{2\pi}} e^{-2(x-1)^2} \mathrm{d}x \rightarrow eqn.3$$

Substituting eqn 2 and 3 in eqn 1, we get,

$$E(penalty) = c.\frac{2}{3} \int_b^\infty \frac{1}{\sqrt{2\pi}} e^{\frac{-(x)^2}{2}} \mathrm{d}x + \frac{1}{3} \int_{-\infty}^b \frac{2}{\sqrt{2\pi}} e^{-2(x-1)^2} \mathrm{d}x$$

We need to minimize the expected penalty, so we can take the derivative of the above equation and set it to zero,

$$\nabla[E(penalty)] = \frac{2c}{3} \left[\frac{(-1)e^{\frac{-b^2}{2}}}{\sqrt{2\pi}} \right] + \frac{1}{3} \left[\frac{2}{\sqrt{2\pi}} e^{-2(b-1)^2} \right] = 0$$

Therefore,

$$c.e^{\frac{-b^2}{2}} = e^{-2(b-1)^2}$$

Taking the natural log of both sides,

$$lnc - \frac{b^2}{2} = -2(b-1)^2$$

$$2lnc - b^2 = -4(b^2 + 1 - 2b)$$

Therefore,

$$3b^2 - 8b + (2lnc + 4) = 0$$

Solving for b,

$$b = \frac{8 \pm \sqrt{64 - (4)(3)(2lnc+4)}}{(2)(3)}$$

$$b = \frac{8 \pm \sqrt{16 - 24lnc}}{6}$$

$$b = \frac{4 \pm \sqrt{4 - 6lnc}}{3}$$

Thus the class conditional densities intersect at two points, but we need to use the lower of the two points as our threshold. (Note: This can be confirmed by plotting the densities as done in the lecture slides.)

Therefore, the new classifier is as below:

$$f^*(x) = \begin{cases} 0 & \text{if } x \leq \frac{4-\sqrt{4-6lnc}}{3} \\ 1 & \text{if } x > \frac{4-\sqrt{4-6lnc}}{3} \end{cases}$$

* * * * * * * * * * * * * * * * *