CS156a Problem Set #5

1. We have

Ein­ ­=  2 ()

And = 0.1, d = 8, and Ein = 0.008

0.008 (0.1)2 (1 - )

N 45

So, the smallest answer is [c].

2. g(x) = sign (w0 (1) + w1(x12) + w2(x22­)) and we can assume that w0can already be selected. Thus, we need to determine w1 and w2 to give us a hyperbola. When we have a circle as our boundary, both of these weight values are negative. This corresponds to the equation of a circle where x12 + x22 = r2. Thus, the values for the hyperbola should be the opposite of what they normally are. Since the hyperbola runs along the y-axis, the equation is x12 / a2 – x22 / b2 = 1. This means that w1 should be negative and w2 should be positive.

w1 < 0 and w2 > 0

The answer is [d].

3. 4(x) = (1, x1, x2, … , x22) and to find the VC dimension, we must count the number size of the tuple:

(z0, z­1, … , z14)

So, d = 14 and

dvc d + 1

dvc 15

So, the answer is [c] since 15 is not smaller than the VC dimension but equal.

4. E(u, v) = (uev– 2ve­-u ­) 2 so differentiate with respect to u

So, using chain rule:

2(uev– 2ve­-u ­)(ev – (-2ve-u))

Thus,

2(uev– 2ve­-u ­)(ev + 2ve-u)

Which matches [e].

5. We need to differentiate in terms of v as well:

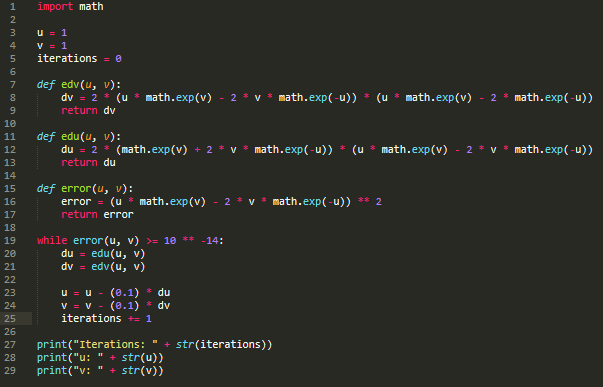
2(uev– 2ve­-u ­)(uev – 2e­-u)

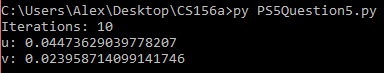
Thus, we create our two functions:

ut+1 = ut – (0.1)

vt+1 = vt – (0.1)

Thus, we can iterate until the error is below 10-14 starting with (u, v) = (1, 1):

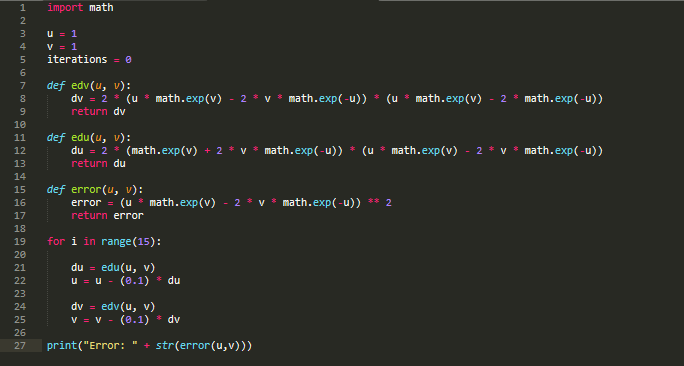




The answer is [d].

6. The closest values are the rounded versions. The answer is [e].

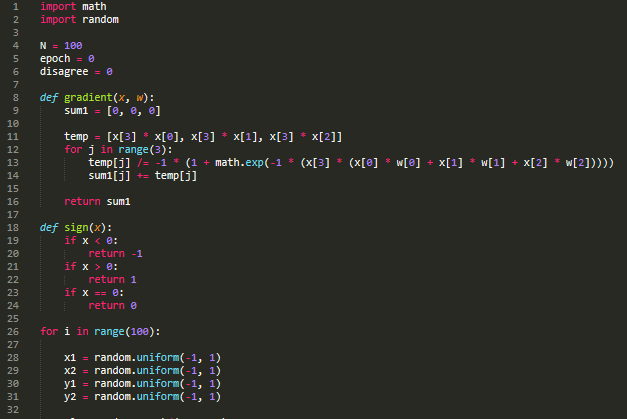
7. Using coordinate descent:

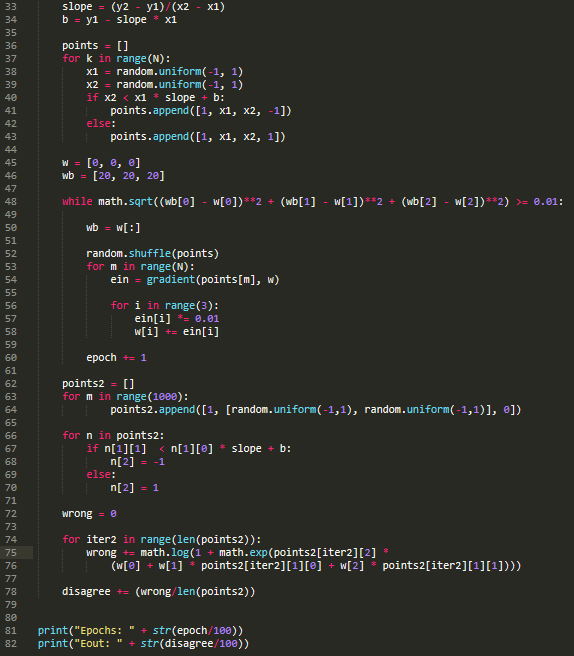




So, the answer is [a].

8. All my work is in my code:







The answer is closest to [d].

9. The answer is closest to [a].

10. For the Perceptron Learning Algorithm, we updated the weight based on if a point was misclassified or not:

wt+1 = wt + ynxn for sign(wTxn) != yn or wt+1 = wt + 0 for sign(wTxn) = yn

If we look at the SGD weight function:

wt+1 = wt - η∇*en ­­*(w)

wt+1 = wt - η(-ynxn / 1 + e ^ ynwTxn)

wt+1 = wt + ynxn ( η / 1 + e ^ ynwTxn)

If we look at this equation, we can see that the expression underlined will either become zero or the ynxn term, which is fully written as yn­wTxn. This corresponds to the same methods of the PLA when η = 1.

The answer is [e].