HW3

December 8, 2023

0.1 CS156A Homework 3

0.2 Wilson Duan

0.2.1 Problem 1.

Since the probability bound is at most 0.03,

$$2Me^{-2\epsilon^2 N} \le 0.03$$
$$2(1)e^{-2(0.05^2)N} \le 0.03 \qquad \text{[Setting } \epsilon = 0.05, M = 1]$$
$$2e^{-0.005N} \le 0.03$$
$$e^{-0.005N} \le 0.015$$
$$-0.005N \le ln(0.015)$$
$$N \ge -200ln(0.015) = 839.941$$

Therefore, the answer is **b)**.

0.2.2 Problem 2.

Repeating for $M = 10$:

$$2Me^{-2\epsilon^2 N} \le 0.03$$
$$2(10)e^{-2(0.05^2)N} \le 0.03 \qquad \text{[Setting } \epsilon = 0.05, M = 10]$$
$$20e^{-0.005N} \le 0.03$$
$$e^{-0.005N} \le 0.0015$$
$$-0.005N \le ln(0.0015)$$
$$N \ge -200ln(0.0015) = 1300.458$$

Therefore, the answer is **c)**.

0.2.3 Problem 3.

Repeating for $M = 100$:

$$2Me^{-2\epsilon^2 N} \le 0.03$$
$$2(100)e^{-2(0.05^2)N} \le 0.03 \qquad \text{[Setting } \epsilon = 0.05, \, M = 100\text{]}$$
$$200e^{-0.005N} \le 0.03$$
$$e^{-0.005N} \le 0.00015$$
$$-0.005N \le ln(0.00015)$$
$$N \ge -200ln(0.00015) = 1760.975$$

Therefore, the answer is **d)**.

0.2.4 Problem 4.

The smallest break point for a perceptron in $\mathbb{R}^3$ is 5 because the break point increases by one as you increase the dimension by one by the VC Dimension Theory. The break point for a perceptron in $\mathbb{R}^2$ is 4, so the break point for a perperceptron in $\mathbb{R}^3$ is $4 + 1 = 5$, and the answer is **b)**.

0.2.5 Problem 5.

If the hypothesis has a break point, its growth function is polynomial. Otherwise, its growth function is 2^N. Looking at each option, i) $1 + N$ is polynomial, so it is a valid growth function.

- ii) $1 + N + \binom{N}{2} = 1 + N + \frac{N(N-1)}{2} = 1 + N + \frac{N^2 - N}{2}$, which is polynomial, so it is a valid growth function.

- iii) $\sum_{i=1}^{\lfloor \sqrt{N} \rfloor} \binom{N}{i}$ is neither polynomial nor 2^N, so it is not a valid growth function.

- iv) $2^{\lfloor N/2 \rfloor}$ is neither polynomial nor 2^N, so it is not a valid growth function.

- v) 2^N is a valid growth function.

Therefore, the answer is **b)**.

0.2.6 Problem 6.

With this hypothesis, there can only be 2 different sets of +1. This hypothesis can shatter 4 points since it is impossible to create 3 different sets of +1. However, with 5 points, you can create 3 different sets of +1 with the points +1, -1, +1, -1, +1, making 5 the break point and the answer to be **c)**.

0.2.7 Problem 7.

There are $\binom{N+1}{4}$ ways to construct two intervals with non-negative length, as you have $N+1$ places to put four "bars", two of which dictate the left bound of each interval, and the other two of which dictate the right bounds of each interval. There are $\binom{N+1}{2}$ ways to construct one interval with non-negative length, as you have $N + 1$ places to put a left and right bound on the interval. There is 1 way to construct no interval, leaving all the points to be classified as -1. As a result, the sum $\binom{N+1}{4} + \binom{N+1}{2} + 1$ is the answer, and the answer is **c)**.

0.2.8 Problem 8.

The smallest break point of an M-intervals hypothesis is the collection of M + 1 intervals of +1s, separated by -1s in between. This is the smallest break point because M-intervals can only generate M intervals to correctly classify the +1s, leaving one of the +1s to be classified incorrectly. Since this break point needs M + 1 intervals of +1s, separated by M -1s, it needs $M + 1 + M = 2M + 1$ points, so the answer is **d)**.

0.2.9 Problem 9.

First, we set up the points in a circle, as this configuration maximizes the number of dichotomies. We note that a triangle easily classifies +1s when they are close together, and starts to struggle when the +1s are separated by -1s inbetween. Since a triangle has three points, it is able to classify 3 separated groups of +1s. If the number of separated groups of +1s is 4 or more, the hypothesis no longer shatters the points. It is only possible to generate 4 separated groups of +1s if you also use at least 4 -1s to separate the +1s. This means 8 is a breaking point, since you can construct the circle with alternating +1s and -1s to create 4 separated groups of +1s. It is possible to shatter 7 because the most amount of separated groups of +1s that can be created is 3 (there are not enough -1s to create 4 separated groups of +1s). Therefore, the answer is **d)**.

0.2.10 Problem 10.

Looking at the concentric circles, we note that it classifies points solely by the distance from the points to the center of the circles (we will denote this distance to be r). In other words, if r is between a and b, it is classified as +1, and classified as -1 otherwise. This method of classifying is synonymous to the 1-interval hypothesis. There are $\binom{N+1}{2}$ ways to construct one interval with non-negative length, as you have $N+1$ places to put a left and right bound on the interval (synonymous to choosing a and b). There is 1 way to construct no interval (synonymous to setting a = b), leaving all the points to be classified as -1. As a result, the sum $\binom{N+1}{2} + 1$ is the answer, and the answer is **b)**.