CS 156a Set 3

Cason Shepard

January 18, 2023

problem 1

Given $M = 1$:

$$\begin{gathered} 0.03 \leq 2(1)e^{-2(0.05)^2 N} \\ ln(0.015) \leq -2(0.05)^2 N \\ 2.099 \geq 0.05^2 N \\ 839.94 \geq N \\ \approx 1000 \end{gathered}$$

The answer is [b]

problem 2

Given $M = 10$:

$$\begin{gathered} 0.03 \leq 2(10)e^{-2(0.05)^2 N} \\ ln(0.0015) \leq -2(0.05)^2 N \\ 3.25 \geq 0.05^2 N \\ 1300 \geq N \\ \approx 1500 \end{gathered}$$

The answer is [c]

problem 3

Given $M = 100$:

$$\begin{gathered} 0.03 \leq 2(100)e^{-2(0.05)^2 N} \\ ln(0.00015) \leq -2(0.05)^2 N \\ 4.4 \geq 0.05^2 N \\ 1761 \geq N \\ \approx 2000 \end{gathered}$$

The answer is [d]

problem 4

We are looking for the least number of points in a 3D space that cannot be shattered by a plane. In the case of 4 points, there is no configuration that is not shattered by a plane, thus $k > 4$. In this case, 5 points will work. If 4 of the 5 points are constructed in a tetrahedron formation, with the final point located somewhere outside, we can guarantee that the outlier point will always be included when we isolate a side of the tetrahedron. Since this point cannot be avoided for the given side, these 5 points are not able to be shattered.

The answer is [b]

problem 5

In Lecture 6, we proved that $m_H(N)$ is polynomial. Thus, the following options must be polynomial for them to be considered valid.

(i) **As shown on Lecture 5 Slide 18, this is valid for positive rays.**

(ii) **This is equivalent to $\frac{1}{2}N^2 + \frac{1}{2}N + 1$, thus, by Lecture 5 Slide 18, is valid for positive intervals**

(iii) By the proof on Lecture 6 Slide 11, the max power of this function can be simplified to $N^{\sqrt{N}}$. This is not polynomial, so it is not valid.

(iv) This function is not polynomial, so it is not valid.

(v) **As shown on Lecture 5 Slide 18, this is valid for convex sets.**

The answer is [b]

problem 6

The smallest break point for this hypothesis set is 5. This is because there is no way to chose 2 intervals of +1 that split up the points as follows: {+1, -1, +1, -1, +1} This is the case that is left out, and is thus shattered. It would require the use of 3 intervals of +1 in order to split up the points like this, which we are unable to do by the specifications of the hypothesis.

The answer is [c]

problem 7

We must consider the arrangement of the intervals. If they are overlapping, the max configurations of the points can be expressed as $\binom{N+1}{2}$. If they are non-overlapping, the max configurations of the points can be expressed as $\binom{N+1}{4}$. The final case (in which none of the points are affected by the intervals selected) results in the additional 1. Thus, the expression becomes:

$$m_H(N) = \binom{N+1}{2} + \binom{N+1}{4} + 1$$

The answer is [c]

problem 8

For each interval in M, we have 2 points. This gives us 2M points that can be correctly classified by M intervals. However, adding one more point makes this impossible to correctly classify. Thus, for M intervals, $2M + 1$ points is the break point.

The answer is [d]

problem 9

In practice, I was getting conflicting answers. I was unable to shatter the 9 point set, but 1 and 3 were easy to shatter.

The answer is [c or d]

problem 10

This problem is similar to the 1-interval problem. We have two points that represent the starting intervals. Using concentric circles, we know that these intervals will be overlapping. Thus, like problem 7, we will have $m_H(N) = \binom{N+1}{2} + 1$.

The answer is [b]