למידה חישובית - ש.ב. 5

שאלה 1

א) נוכיח כי $VC(H)=\infty$ ע"י ניפוץ כל דיכוטומיה שקיימת במימד
כאשר $A\in\{-1,1\}$, וזה כל קב' מוכלת ב-X.

$$F(x)=\begin{cases}-1 & x\in S\\ 1 & \text{else}\end{cases}$$

הפונקציה מנפצת כל קבוצה בכל אורך וקיים מיפוי לכל קב'.

ב) נוכיח כי $VC(H)=2n$ כאשר n הוא כמות האינטרוולים

$VC(H)\geq 2n$:

עבור 2n איברים נבחר את הData שהמרחקים ביניהם הוא 2 (בין כל אחד לשני).

כאשר הData הומוגני: נגדיר שהאינטרוול יהיה לפני הנק' הראשונה ויסתיים אחרי הנק' האחרונה.
שאר האינטרוולים נגדיר בגודל 2 אחרי הנק' האחרונה.

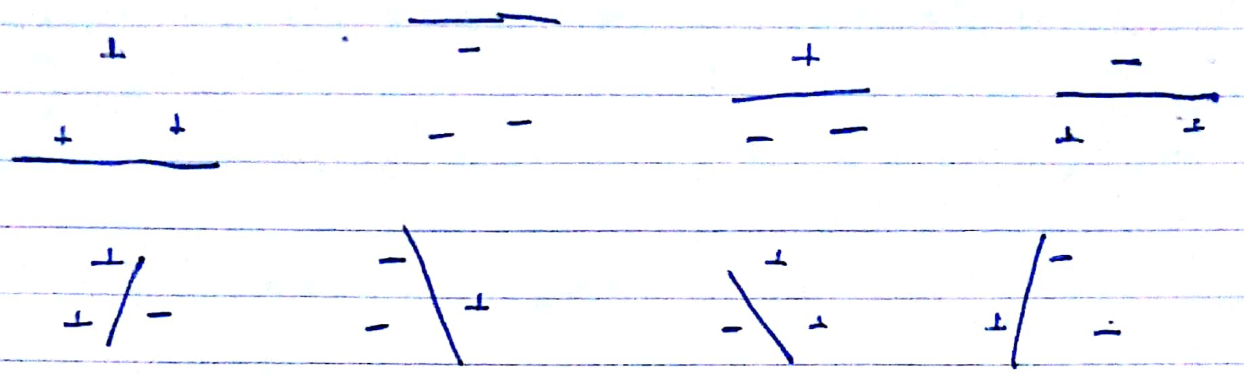
כאשר הפריסה היא הנק' הן "+" לסירוגין "-": יש n נק' שאמורות "+" ויש n נק' שאמורות "-"
נבחר בצורה n אינטרוולים כך שכל אינטרוול מכיל נק' 1 "+" בלבד.

כל אפשרות אחרת: יהיו לפחות 2 נק' בעלי אותו סימן ונאגד אותם תחת אינטרוול.

נראה כי $VC(H)\ngeq 2n+1$:

נסתכל עבור הפריסה $+-+-+\dots-+$ יש n "-" ו n+1 "+"
אבל יש רק n אינטרוולים ולכן אי אפשר להפריד את כל ה n+1 אחד.

ג) נראה חלוקה של ג מופעים שמתנפצים ע"י הדיכוטומיות:



ד) נראה עבור y אז יתקיים $h((1-\alpha)z+\alpha z')=1$ $\forall\alpha\in[0,1]$

$$h((1-\alpha)z+\alpha z') = h((1-\alpha)x_1+\alpha x_1',\ (1-\alpha)x_2+\alpha x_2')$$

$$= w_1(x_1-\alpha x_1+\alpha x_1') + w_2(x_2-\alpha x_2+\alpha x_2') + b =$$

$$w_1x_1 - \alpha x_1 w_1 + \alpha w_1 x_1' + w_2x_2 - w_2\alpha x_2 + \alpha w_2 x_2' + b =$$

$$\underline{w_1x_1+w_2x_2} - \alpha(w_1x_1+w_2x_2) + \alpha(w_1x_1'+w_2x_2') + \underline{b} + \alpha b - \alpha b =$$

$$= h(z) - \alpha h(z) + \alpha h(z') = h(z) = y$$

<u>עבור חלוקת זוגית</u>: אזי המקרה: $- + - +$ לא ניתן להפריד ע"י קו לינארי את ה-"ים מהאחרים.

<u>עבור חלוקת משולש</u>: $- \overset{+}{-}$ לא ניתן להפריד משאר המינוסים כצורת מפריד לינארי (זהו בעצם xor).

<u>עבור חלוקת מרובע</u>: $\begin{matrix} + & - \\ - & + \end{matrix}$ זהו בעצם xor אשר הוכחנו בהרצאה/תרגול שלא ניתן להפריד לינארית.

שאלה 2

נגדיר את ההיפותזה: $h=(x_1\wedge\bar{x}_1)\wedge\ldots\wedge(x_n\wedge\bar{x}_n)$

ולכל $1\le i\le n$:

עבור h: אם $c(x_i)=1$ אז מחק את $\bar{x}_i$

אם $c(\bar{x}_i)=0$ אז מחק את x_i.

האלגוריתם עקבי עם המידע כי הוא נבנה בהתבסס על הקונספט.

סיבוכיות זמן ריצה:

ישנן 3 אפשרויות לכל תכונה ולכן $|H|=3^n$:

$$m\ge \frac{1}{\varepsilon}\left(\ln(|H|)+\ln\left(\frac{1}{\delta}\right)\right) = \frac{n}{\varepsilon}\ln(3)+\frac{1}{\varepsilon}\ln\left(\frac{1}{\delta}\right) =$$

ולכן הוא זמן פולינומיאלי.

ומכאן האלגוריתם הינו PAC-learnable.

שאלה 3

נגדיר אלגוריתם A:

לכל x במידע:

$$h(x) = \begin{cases} 1 & c(x)=1 \text{ אם} \\ L(x) & \text{אחרת} \end{cases}$$

סיבוכיות זמן ריצה:

אנו יודעים כי L כל הזמן פולינומיאלי. לכן אם נבחר $L(x)$ אז PAC-learnable, בסיבוכיות $m(\delta,\varepsilon)$.

אם נבחר ב-1 אזי $|H|=1$ ולכן נקבל:

$$m \geq \frac{1}{\varepsilon}\left(\ln(H) + \ln\left(\frac{1}{\delta}\right)\right) = \frac{1}{\varepsilon}\left(\ln(1) + \ln\left(\frac{1}{\delta}\right)\right) = \frac{\ln\frac{1}{\delta}}{\varepsilon}$$

לכן סיבוכיות זמן הריצה הנו: $\max\left\{m(\delta,\varepsilon), \left\lceil \frac{\ln\frac{1}{\delta}}{\varepsilon} \right\rceil\right\}$

Scanned by CamScanner