IML - תרגיל 2

1)

$$Im(A) = \{x \in V \mid \exists v \in V \; Av = x\}$$

$$ker(A) = \{A(v) = 0 \mid v \in V\}$$

שקילות: $ker(A^T)^\perp = \{x \in R^d \mid \langle x, v\rangle = 0 \;\; \forall v \in V \;\; A^T v = 0\} =$

$ker A^T = ker(A^T A)$ $\quad x \in R^d \mid \langle x, v\rangle = 0 \;\; \forall v \in V \;\; A^T A v = 0\}$

(נראה הכלה דו כיוונית:

יהי $Aw = x$, $v \in V$ כך ש- $A^T A v = 0 \Leftrightarrow A^T v = 0$

$$\langle x | v\rangle = x^T v = (Aw)^T v = w^T A^T v = 0$$

לכן $x \in ker(A^T)^\perp$

כיוון שני:

$$\langle x | v\rangle = 0 = x^T v \;\wedge\; A^T v = 0 \Leftrightarrow A^T A v = 0$$

נניח בשלילה כי לא קיים $v \in V$ כך ש- $Av = x$ לכן כי הכיוון של x, ואז y מאונך כך ש- $\langle A | y\rangle = 0$ לכן הגענו לסתירה לכך ש- $A^T y = 0$ לכל $y \in V$ ולכן $\langle x | y\rangle = 0$ כלומר, $y \perp x$ בסתירה לכך ש- x לא ב- ImA, $x \in ImA$.

2) הנלכם $ker A$, $Ax = b$.

$$ker(A^T) = ker(AA^T) = \{v \in R^d \mid AA^T(v) = 0\}$$

יהי $v \in ker A^T$, $\langle b | v\rangle = b^T v = 0$.

אם A לא הפיכה אז יש לה אינסוף פתרונות או אין פתרונות,

הוכחה: בגלל שנתון Gram-שמידט, כי אם $b \in ImA$ אז יש פתרון יחיד.

אם $b \perp ker A^T$ אז $b \in ker(A^T)^\perp$ וזה סותר A לא הפיכה, לכן:

$b \in ImA \Longleftrightarrow$ יש פתרון יחיד $\Leftrightarrow$ יש אינסוף פתרונות כי A

אינה הפיכה ולכן פתרון יחיד.

3) נעבור למקרה הזה - אם נסמן $xx^T = A$, $xy = b$, ואז הגדרת המשוואה $Aw = b$

אם נוכל 2 אם $A = xx^T$ אז הטענה זהה אם $xy \perp ker(xx^T)^T$ יש אינסוף

פתרונות. $ker((xx^T)^T) = ker(xx^T)$.

$$ker(xx^T) = \{z \in R^d \mid xx^T z = 0\} = \underset{\Downarrow}{\{z \in R^d \mid x^T z = 0\}}$$

$$ker\, x^T = ker(xx^T)$$

$$\langle xy | z\rangle = y^T \underset{\Downarrow}{x^T} = 0$$

$$x^T z = 0$$

ואז קיבלנו כי $xy \perp ker(xx^T)^\perp$ לכן יש xx^T לא הפיך, אזי יש אינסוף

פתרונות. אם xx^T הפיכה אז משפט 1 ולכן יחיד כי אם המשוואה

שלי המלאה יש פתרון יחיד.

4.a) נשים לב כי $X = U\Sigma V^T$.

$$X^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

$$XX^T = U\Sigma V^T V\Sigma^T U^T =$$

כי V אורתוגונלי ולכן $V^TV=I$

$$U\Sigma\Sigma^T U^T = UDU^T$$

$$D = \Sigma\Sigma^T$$

כיוון ש- U, U^T אורתוגונליים ולכן הפיכות כי

$$(XX^T)^{-1} = UD^{-1}U^T$$

$$(XX^T)^{-1}X = UD^{-1}U^TU\Sigma V^T = UD^{-1}\Sigma V^T =$$

$$U\begin{pmatrix} a_1^{-2} & & & \\ & \ddots & & \\ & & a_r^{-2} & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}\begin{pmatrix} a_1 & & & \\ & \ddots & & \\ & & a_r & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix}V^T = U\Sigma^{T^{-1}}V^T =$$

$D = \Sigma\Sigma^T$

שערכי

והאיברים באלכסון ב- Σ, Σ^{-1} הם הפיכים!

ההגדרה של Σ^T

$$= U(\Sigma^T)^+ V^T$$

$$X^{T+} = (X^+)^T = (U\Sigma^+V^T)^T = U(\Sigma^+)^T V^T = \underline{\underline{U(\Sigma)^T V^T}}$$

$$(X^+)^+ = (XX^T)^{-1}X$$

מש"ל

b) $\{x_1 \ldots x_m\} = \mathbb{R}^d \Longleftrightarrow XX^T$ הפיך

$$XX^T = \begin{bmatrix} x_1 & \ldots & x_m \end{bmatrix}\begin{bmatrix} x_1^T \\ \vdots \\ x_m^T \end{bmatrix} = \begin{bmatrix} \sum x_i^1 x_i^1 & \ldots & \sum x_i^1 x_i^m \\ \sum x_i^1 x_i^m & \ldots & \sum x_i^m x_i^m \end{bmatrix}$$

אם XX^T הפיך $\Longleftrightarrow 0, \dim(\ker(X^TX)) = \dim(\ker(X))$ ומכאן נובע ש:

$\Longleftrightarrow \dim(Im(XX^T)) = \dim(Im(X)) = d$, מכאן ש- XX^T בגודל $d \times d$ ולכן בהכרח ה-svd

של Σ בגודל $d \times d$ ולכן מהצורה הזו $\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_d \end{pmatrix}$, ומהגדרה יש לו d ערכים סינגולריים ולכן

יש לו עמודות בת"ל d (לפחות) $\Longleftrightarrow \{x_1 \ldots x_m\} = \mathbb{R}^d$.

כי אם יש לו עמודות X בת"ל של המטריצה X מכסות, ולכן נוכל

XX^T הפיך $\Longleftrightarrow$ ש- $\{x_1 \ldots x_m\} = \mathbb{R}^d$.

$$XX^Tw = Xy$$
$$UDU^Tw = U\Sigma V^Ty$$
$$DU^Tw = \Sigma V^Ty$$
$$\begin{pmatrix} a_1^2 & & \\ & \ddots & \\ & & a_r^2 \\ & & & \ddots \end{pmatrix} U^Tw = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_r \\ & & & \ddots \end{pmatrix} V^Ty$$

ולכן נגדיר עבור ה- $i \geq r$ שיהיה שווה 0, (ונכפיל) בצדדים את σ^+:

$$\boxed{\sigma^+ DU^Tw = \sigma^+\Sigma V^Ty}$$

$$\begin{pmatrix} a_1^{-2} & & \\ & \ddots & \\ & & a_r^{-2} \end{pmatrix}\begin{pmatrix} a_1^2 & & \\ & \ddots & \\ & & a_r^2 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 0 \\ & & & & 0 \end{pmatrix}$$

$$\begin{pmatrix} a_1^{-2} & & \\ & \ddots & \\ & & a_r^{-2} \\ & & & 0 \\ & & & & 0 \end{pmatrix}\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_r \end{pmatrix} = \begin{pmatrix} a_1^{-1} & & \\ & \ddots & \\ & & a_r^{-1} \end{pmatrix} = \Sigma^+$$

$U \,|\, U^Tw = \Sigma^+ V^T y$ נכפיל ב:

$$w = U\Sigma^+ V^T y$$

$$\boxed{w = X^{+T} y}$$

בהגדרה הזאת $w_i = 0$ לכל $r+1 \leq i$ לא משנה מה נבחר עדיין יהיה בגודל סופי,
ולכן $\bar{w}$ יהיה זהה ל- $\hat{w}$ על קורדינטות $r \geq i$ שווה ולכן עבורם הם
$\bar{w}_i$ ו- a_i הם מוגדרים, ועבור כל $r+1 \leq i$ הערך a ו- w_i לא משנה כי הוא שווה
$\bar{w}_i = \hat{w}_i$ $r \geq i$ (עבור) לכן נראה כי:
נראה עכשיו כי הנורמה של $\hat{w}$: $|\hat{w}|^2 = \langle \hat{w} | \hat{w} \rangle = (\hat{w})^T \hat{w} = \sum_{i=1}^{d} \hat{w}_i \hat{w}_i = \sum_{i=1}^{r} \hat{w}_i \hat{w}_i$

ועכשיו של $\bar{w}$ נקבל: $\|\bar{w}\|^2 = \langle \bar{w} | \bar{w} \rangle = \sum_{i=1}^{d} \bar{w}_i \cdot \bar{w}_i = \sum_{i=1}^{r} \hat{w}_i \cdot \hat{w}_i + \sum_{i=r+1}^{d} \bar{w}_i \cdot \bar{w}_i$

ומכיוון כי מתקיים $\sum_{i=r+1}^{d} \bar{w}_i \cdot \bar{w}_i \geq 0$ ולכן $\|\bar{w}\|^2 \geq \|\hat{w}\|$ ומכאן, $\hat{w}$ הוא
בעל הנורמה המינימלית. ∎

4.a) נשים לב כי $X = U\Sigma V^T$.

$$X^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

$$XX^T = U\Sigma V^T V \Sigma^T U^T =$$

$$U\Sigma\Sigma^T U^T = UDU^T$$ כי V אורתוגונלית ולכן $V^TV = I$

$$D = \Sigma\Sigma^T$$

$$(XX^T)^{-1} = UD^{-1}U^T$$ כיוון ש-U, U^T אורתוגונליות ולכן הפיכות

$$(XX^T)^{-1}X = UD^{-1}U^TU\Sigma V^T = UD^{-1}\Sigma V^T =$$

$$U\begin{pmatrix} a_1^{-2} & & & \\ & \ddots & & \\ & & a_r^{-2} & \\ & & & 0_{0_0} \end{pmatrix}\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_r \\ & & & 0_{0_0} \end{pmatrix}V^T = U\Sigma^{T^{-1}}V^T =$$

$D = \Sigma\Sigma^T$

שערכי האלכסון ב-Σ, Σ' הם הפוכים!

$$\Sigma^T = U(\Sigma^{T_1})^+ V^T$$ הגדרה ש

$$X^{T+} = (X^+)^T = (U\Sigma^+ U^T)^T = U(\Sigma^+)^T V^T = \underline{U(\Sigma^T)^+ V^T}$$

$$(X^T)^+ = (XX^T)^{-1}X$$ ומכאן

b) XX^T הפיך $\Leftrightarrow \{x_1 \dots x_m\} = \mathbb{R}^d$.

$$XX^T = \begin{bmatrix} x_1 & \dots & x_m \end{bmatrix}\begin{bmatrix} x_1^T \\ \vdots \\ x_m^T \end{bmatrix} = \begin{bmatrix} \sum x_i^1 x_i^1 & \dots & \sum x_i^1 x_i^m \\ \sum x_i^1 x_i^m & \dots & \sum x_i^m x_i^m \end{bmatrix}$$

אם XX^T הפיך $\Leftrightarrow 0 = \dim(\operatorname{Ker}(X^TX)) = \dim(\operatorname{Ker}(X))$ ומכאן נובע ש:
$d = \dim(\operatorname{Im}(XX^T)) = \dim(\operatorname{Im}(X))$, ומכאן ש-$XX^T$ מגודל $d \times d$ ולכן בדירוג ה-svd
של Σ בגודל $d \times d$ ולכן מופיעה בה $(a_1 \dots a_d)$, אלכסונית, לפחות d ערכים סינגולריים ולכן
מימד התמונה הוא d (למשל) $\Leftrightarrow \{x_1 \dots x_m\} = \mathbb{R}^d$.
וגם המימד של X זה המימד של עמודות X כלומר, ולכן מיקום
כי XX^T הפיך $\Leftrightarrow$ span$\{x_1 \dots x_m\} = \mathbb{R}^d$.

ⓒ 4

$$XX^Tw = Xy$$
$$UDU^Tw = U\Sigma V^Ty$$
$$DU^Tw = \Sigma V^Ty$$
$$\begin{pmatrix} a_1^2 & & \\ & \ddots & \\ & & a_r^2 \\ & & & \ddots \end{pmatrix} U^Tw = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_r \\ & & & \ddots \end{pmatrix} V^Ty$$

ולכן נגדיר עבור פיצוי $i \geq r$ שיהיה שווה 0, נרצה לבודד את σ^+:

$$\boxed{\sigma^+ \sigma U^Tw = \sigma^+ \Sigma V^Ty}$$

$$\begin{pmatrix} a_1^{-2} & & \\ & \ddots & \\ & & a_r^{-2} \end{pmatrix}\begin{pmatrix} a_1^2 & & \\ & \ddots & \\ & & a_r^2 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \\ & & & & 0 \end{pmatrix}$$

$$\begin{pmatrix} a_1^{-2} & & & \\ & \ddots & & \\ & & a_r^{-2} & \\ & & & 0 \\ & & & & 0 \end{pmatrix}\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_r \end{pmatrix} = \begin{pmatrix} a_1^{-1} & & \\ & \ddots & \\ & & a_r^{-1} \end{pmatrix} = \Sigma^+$$

סיכום לא:

$$U\cdot| \quad U^Tw = \Sigma^+ V^Ty$$
$$w = U\Sigma^+ V^Ty$$
$$\boxed{w = X^{+T}y}$$

בהגדרה הזאת $w_i = 0$ לכל $i > r$ כלומר שפתרון זה ... ואם בקישור נוסף,
נחפש $\bar{w}$ אחר לפיה $\hat{w}$ של קורדינאטות $w_i \leq r$ נותנות ואלו שונות הן
a_i ו- $\hat{w}_i$ ... , חוץ מעבור פיצוי a ... לא תתקיים כי ...
$\bar{w}$ ... $\hat{w}$ ... $\bar{w}_i = \hat{w}_i$ ... 4 נראה ש:

$$|\hat{w}|^2 = \langle \hat{w} | \hat{w} \rangle = (\hat{w})^T \hat{w} = \sum_{i=1}^{r} \hat{w}_i \cdot \hat{w}_i = \sum_{i=1}^{r} \hat{w}_i \hat{w}_i$$

ועבור $\bar{w}$ נקבל:

$$\|\bar{w}\|^2 = \langle \bar{w} | \bar{w} \rangle = \sum_{i=1}^{d} \bar{w}_i \cdot \bar{w}_i = \sum_{i=1}^{r} \hat{w}_i \cdot \hat{w}_i + \sum_{i=r+1}^{d} \bar{w}_i \cdot \bar{w}_i$$

ונשים לב כי מתקיים $\sum_{i=r+1}^{d} \bar{w}_i \cdot \bar{w}_i \geq 0$ ולכן $\|\bar{w}\|^2 \geq \|\hat{w}\|$ ומכאן, $\hat{w}$ הוא
בעל הנורמה המינימלית. ∎

9. נרצה להראות כי $VCdim = n$.

אחילה נראה כי $n \le VCdim$

ניקח את הבסיס הסטנדרטי $X' = \{e_1, \dots, e_n\}$, $C \subset X'$, $y = (y_1, \dots, y_n)$

המתאימים ל-X ונראה שקיבלנו C. נגדיר $Y_1 = \{i \mid y_i = 1\}$, $Y_0 = \{i \mid y_i = 0\}$

עבור ה-y שקיבלנו, נגדיר את הקבוצה I להיות $I = \{i \mid y_i \in Y_1\}$

נשים לב כי עבור e_j ש-$y_j = 1$.

מתקיים כי $\left(\sum_{i \in I} x_i\right) \bmod 2 = 1$ (כיון ש-$j \in I$) כלומר $h_j = y_j$,

עבור e_j ש-$y_j = 0$. $\sum_{i \in I} x_i \bmod 2 = 0 \bmod 2 = 0$ — כיון שאין ש-$j \in I$ ולכן $h_j = y_j$

ונראה כי $VCdim \le n$. נסתכל על הקבוצה $\{x_1, \dots, x_{n+1}\}$ ונראה

להראות כי לכל תיוג של y: $\{y_1, \dots, y_{n+1}\}$ ש-$h \ne y$.

נשים לב כי כיון שהוקטורים בממד n יש להם תלות לינארית ולכן $x_{n+1} = \sum_{i=1}^{n} b_i x_i$

נסמן $y_i = \operatorname{sign} b_i$ עבור $1 \le i \le d$ ו-x_i עם התיוג הקבוצה מקבלת את x_{n+1}.

$$h(x_{n+1}) = \left(\sum_{i \in I} x_i\right) \bmod 2 = \left(\sum_{i \in I} \sum_{j=1} b_j x_i^j\right) \bmod 2 =$$

$$\left(\sum_{j=1}^{d} \sum_{i \in I} b_j x_i^j\right) \bmod 2 = \left(\sum_{j=1}^{d} \sum_{i \in I} b_j x_i^j\right) \bmod 2 = \sum_{j=1}^{d} b_j \left(\left(\sum_{i \in I} x_i^j\right) \bmod 2\right) \bmod 2 =$$

$$\left(\sum_{j=1}^{d} b_j \operatorname{sign} b_j\right) \bmod 2 = \left(\sum_{j=1}^{d} b_j \operatorname{sign} b_j\right) \bmod 2 = \sum_{j=1}^{d} |b_j| \bmod 2.$$

ולכן אם ניקח $y_{n+1} \ne \sum_{i=1}^{d} (b_i |^{-1}) \bmod 2$ נקבל את התיוג ההפוך למה

שיוצא לנו, כלומר לא ניתן לקבל כזה y_{n+1} ולכן קיבלנו כי $n \ge VCdim$

10. יהי k מספר טבעי ו-$([a_i, b_i])_{i=1}^{k}$ סדרה של k אינטרוולים ב-R.

$A = \bigcup_{i=1}^{k} [a_i, b_i]$, $H_{k\text{-intervals}}$ כולל את $h_A(x) = \begin{cases} 0 & x \notin A \\ 1 & x \in A \end{cases}$ עבור כל b ו-a כאלה ו-k אינטרוולים.

נראה ש-$VCdim = 2k$.

נראה אחילה כי $VCdim < 2k+1$.

יהי $x_1, \dots, x_{2k+1}$ נקודות. נגדיר את התיוג: $y_i = j \bmod 2$ כלומר התיוג שמתחלף

הוא מהצורה $y = \{1010 \dots 101\}$, נספור מספר ה-1 ב-k אינטרוולים כדי

לכסות מצורך המרחק 2 נקודות שיהיו באותו מקטע. אם אז לא ניתן לכסות

מכך כן יש אם שמכסים בין 2 אחדות כלן אך על לפחות $2k+1$ נקודות.

כעת, נראה שניתן לקבל אם כן, מאחר המקטע השמאלי ביותר שבמקטע את ניתן לכסות

2 נקודות, ולכן כאשר נקח את k מקטעים ומכסים כי ניתן לכסות

k נקודות, כנדרש.

ב) הראו שאם A אלגוריתם סופי ש... (ועקבי) אז H אם ∞-מימד VC.

אם ניקח את כל ... אינסוף 2k נקודות. $\infty = {}$ ... נקודות.

בפרט, נוכל לקחת א"ם ... גדול כרצוננו

חיבור כנה: ∞-מימד VC $\to$ $|C| = \infty$

11. נראה שיש דוגמה גודל $d+1$ מנותצת:

נסמן ב- $C=\{e_1,\dots,e_d,0\}$, $y=(y_1,\dots,y_{d+1})$ ... המקרה שקודם במרצה:

כאשר $y_{d+1}=0$ ניקח $b=0$ ואז לכל ... $w=(y_1,\dots,y_d,0)$ ... $\langle w, e_i\rangle = y_i$ ולכן C

עבור h_w, כלומר נגדיר ...

מנותצת, כאשר $y_{d+1}=1$ ניקח $b=1$ ואז עבור x_{d+1} נקבל:

$$h_{w,b}(x_{d+1}) = \text{sign}(\langle w | x_{d+1}\rangle + 1) = 1 = y_{d+1}.$$

כמו כן לכל $1\le i\le d$ נגדיר $w=(y_1-1,\dots,y_d-1)$, ונקבל

$$h_{w,b}(x_i) = \text{sign}(\langle w|x_i\rangle+1) = \text{sign}(\langle w|e_i\rangle+1) = \text{sign}(y_i-1+1) = \text{sign}(y_i) = y_i$$

כנדרש!

2) נראה כעת שכל קבוצה בגודל $d+2$ לא יכולה להיות מנותצת. נניח בשלילה שקיימת קבוצה בגודל $d+2$ $C=\{x_1,\dots,x_{d+2}\}$ מנותצת. לכן, עבור $(y_1,\dots,y_{d+2})$ קיים וקטור $w\in\mathbb{R}^d$ ו- $b\in\mathbb{R}$ כך ש- $\text{sign}(\langle w|x_i\rangle+b)=y_i$ לכל i.

נגדיר $w'=(w_1,\dots,w_d,b)$ ולכל $1\le i\le d+2$ $x'_i=(x_{i1},\dots,x_{id},1)$

שים לב כי לכל $1\le i\le d+2$ מתקיים: $\text{sign}(\langle w|x_i\rangle+b)=\text{sign}(\langle w'|x'_i\rangle)$

הקבוצה $C'=(x'_1,\dots,x'_{d+2})$ מנותצת ע"י מחלקת המפרידים ההומוגניים ב- $\mathbb{R}^{d+1}$!

בסתירה לכך כי מימד VC שלהם כפי שהוכחנו הוא $d+1$.

Agnostic PAC

13. תהי H מחלקת היפותזות.

א) תהי H למידה PAC, יהי A אלגוריתם למידה עבור H ו-m_H מהווה Sample Complexity מתאים.

נראה ש-H למידה PAC.

יהי D התפלגות על X ו-f פונקציה כלשהי.

נגדיר התפלגות חדשה D' על $\{0,1\} \times X$ עבור הדגימות המתאימות (x,y) באופן הבא: x מוגרל לפי D, ו-y נקבע ע"י f. כך שנקבל ש"מ ($\min_{h \in H} L_{D'}(h)=0$) מתקיים: $\min_{h\in H} L_{D'}(h)=0$

יהי $\epsilon, \delta \in (0,1)$ ו-D' עבור $m \geq m_H(\epsilon,\delta)$

אם נקח m דגימות שמתויגות ע"י f, האלגוריתם A יחזיר בסיכוי של לפחות $1-\delta$ היפותזה h שמקיימת $L_D(h) \leq 0+\epsilon = \epsilon$

$\underset{h\in H}{\min} L_{D'}(h)$.

ב) אם H למידה Agnostic PAC מוכיחים שגם H למידה PAC מוגדרת. נשים לב שהוכחנו קודם שאם H למידה ERM אזי Agnostic PAC ולמעשה הוכחנו ERM עבור PAC, ולכן כי קיבלנו למידה PAC מוגדרת כי קיבלנו h שמקיימת את המינימום הנ"ל.

MONOTONICITY

א. H מחלקת היפותזות עבור סיווג בינארי. נניח ש-H למידה PAC, ותהי m_H, הוכח ש-m_H אלה חלק עבור δ בנפרד והפרמטרים ε, δ.

בהינתן $\delta\in(0,1)$ ו- $0<\varepsilon_1\le\varepsilon_2<1$ מתקיים $m_H(\varepsilon_1,\delta)\ge m_H(\varepsilon_2,\delta)$.

מהגדרה, עבור ε_1 $P\{L_{D,f}(A(S))<\varepsilon\}\ge 1-\delta$.

מתקיים $m\ge m_H(\varepsilon_1,\delta)$.

באופן דומה עבור ε_2 $P\{L_{D,f}(A(S))<\varepsilon_2\}\ge 1-\delta$. וזה נכון $m\ge m_H(\varepsilon_2,\delta)$.

עבור δ $m\in\mathbb{N}$ שמקיים $L_{D,f}(A(S))\le\varepsilon_1$ אזי $L_{D,f}(A(S))<\varepsilon_2$

נבחר כן ש- $m=m_H(\varepsilon,\delta)$ מקיים. ולכן אז מתקיים $L_{D,f}(A(S))<\varepsilon_2$.

ומכאן - $m_H(\varepsilon_1,\delta)\ge m_H(\varepsilon_2,\delta)$

בהינתן $\varepsilon\in(0,1)$ ו- $0<\delta_1\le\delta_2<1$ מתקיים $m_H(\varepsilon,\delta_1)\ge m_H(\varepsilon,\delta_2)$

עבור δ: כי $m\in\mathbb{N}$ שמקיים $P\{L_{D,f}(A(S))<\varepsilon\}\ge 1-\delta_1$ קיים אז

$P\{L_{D,f}(A(S))<\varepsilon\}\ge 1-\delta_2$ כי $\delta_1\le\delta_2$ ו- $1-\delta_1\ge 1-\delta_2$

נבחר, $m=m_H(\varepsilon,\delta_1)$ מקיים את 2 התנאים ולכן $m_H(\varepsilon,\delta_1)\ge m_H(\varepsilon,\delta_2)$

ב. $VCdim$ ש מחלקה מוגדרת בגודל הבא:

$$VCdim(H)=\sup\{|C| \mid h \text{ shatters } C\}$$

אם $C\subseteq X$ מהגדרה ש המקיימת H לכל עבור $H_1\subseteq H_2$ (H_2

המחלקה H מנפצת קבוצה סופית C אם $|H_C|=2^{|C|}$ לכן

C ש-H_{1C} ואינה מנוצלת ב-H_{2C} מוגדר.

לכן קבוצה ש-H_{1C} מנפצת גם H_{2C}

$VCdim(H_1)\le VCdim(H_2)$

$\blacksquare$

Practical:

5. לכל מיקוד עמדי עמודה משני צד הנדסי. להתאים לפי המיקום.
תיקון 4, הורדו עמודות, כאן יש ערכים ריקים.
6. תכונות תשובה - מס' מדדי השינה, מספר מדדי השורות.
תכונות כמות תשובה - צ"כ קוד וקו' רוחב.