

מבוא ללמידה עבודה 3

1. b.

K	lambda	Error rate
3	0.01	0.0740
3	0.1	0.0780
3	1	0.3950
6	0.01	0.0710
6	0.1	0.0770
6	1	0.4400
10	0.01	0.0690
10	0.1	0.0710
10	1	

The optimal pair is $(K, \lambda) = (10, 0.01)$.

The error rate for $(10, 0.01)$ for the entire test set is 0.0700.

1. c.

define psi as:

```
psi = @(x)
[1, x(1), x(2), x(1)*x(2), x(1)^2, x(2)^2, x(1)^2*x(2), x(1)*x(2)^2, x(1)^3, x(2)^3];
```

By using the formula at lecture 8 slide 11 (**Representer Theorem**), we achieve the separator:

$w = [-0.0073, 0.4256, 0.3919, 0.0220, 0.0324, 0.0038, 0.0164, -0.0412, 0.0210, 0.0201]$

The resulting multivariate polynomial is:

$$\begin{aligned} & -0.0073 + 0.4256x_1 + 0.3919x_2 + 0.0220x_1x_2 + 0.0324x_1^2 + 0.0038x_2^2 \\ & + 0.0164x_1^2x_2 - 0.0412x_1x_2^2 + 0.0210x_1^3 + 0.0201x_2^3 \end{aligned}$$

② חוקר מדעיין אולם את אחוז הזמן בו חיצו ברטומה, בסדר α ב-2015

נסמן אחוז זה ב- α . אם כן הוא מאריך נקודת ב-2015

באלון אחי ואקראי בוקר את ארץ הדרום כך ארץ האם סדרה ברטומה

בסדר α האם זמן. גובא-ההקדמה הן $[0, \hat{p}]$ הים בן א הקדמה

אכן סדרה ברטומה אלא הקדמה

(א) נסגרים האי מיליון Hoeffding בן אנסין \hat{p} מסומה ארץ α :

(י) נסביר א \hat{p} מסומה $1 \leq \hat{p} \leq 1$ [קצמן א סדרה ברטומה בסדר α $Z_i = \mathbb{I}[Z_i = 1]$]

100 מסגרים מקריים

(II) התוצאה עבור \hat{p} זמן α ב-2015 קבוצה מרדס כיוון שזה צורה סדרה דור

וא כמעט א הערכו באקראי ובהרבה ארץ על בני 2015 ואן

אן גלוי בן א א קסמה.

(III) (עבור α מיוצג האמזיס) ההסתברות אחרת באקראי נקודת זמן

ב-2015 בה חוקרה ברטומה קוא $P[Z_i = 1] = \frac{\alpha}{100}$ מסומה \forall

(ב) נניח כי $Z_i = \hat{p} - \frac{1}{100} \sum_{i=1}^{100} Z_i$ א \hat{p} א מיליון Hoeffding מתקיים:

$$P\left[\left|\frac{1}{100} \sum_{i=1}^{100} Z_i - \hat{p}\right| \geq \epsilon\right] \leq 2e^{-2 \cdot 100 \cdot \epsilon^2} \quad p = \alpha$$

$$P[|0.22 - \alpha| \geq \epsilon] \leq 2e^{-200 \epsilon^2}$$

$$1 - P[|0.22 - \alpha| < \epsilon] \leq 2e^{-200 \epsilon^2}$$

$$P[0.22 - \epsilon < \alpha < 0.22 + \epsilon] > 1 - 2e^{-200 \epsilon^2}$$

אנו מסומה ההסתברות ארץ 99 אן נקודת:

$$2e^{-200 \epsilon^2} = 0.01 \Leftrightarrow e^{200 \epsilon^2} = 200 \Leftrightarrow \epsilon^2 = \frac{\ln(200)}{200}$$

$$\Leftrightarrow \epsilon^2 = \frac{\ln(200)}{200} \Leftrightarrow \epsilon = \pm \sqrt{\frac{\ln(200)}{200}} \Rightarrow \epsilon = 0.1628$$

$$P[0.22 - 0.1628 < \alpha < 0.22 + 0.1628] > 1 - 2e^{-200(0.1628)^2}$$

נצבי האזרה קוסמה וקראי.

$$P[0.0572 < \alpha < 0.3828] > 0.99$$

כאמר ההסתברות אחרת ארץ 99 מתקיים $0.0572 < \alpha < 0.3828$

$$(3) \text{ k) } f(\omega) = \lambda \|\omega\| + \sum_{i=1}^m (\langle \omega, x_i \rangle - y_i)^2$$

$$(\lambda \|\omega\|)'_{\omega_j} = (\lambda \sqrt{\omega_1^2 + \omega_2^2 + \dots + \omega_d^2})'_{\omega_j} = \frac{\lambda \cdot 2\omega_j}{2\sqrt{\omega_1^2 + \dots + \omega_d^2}} = \frac{\lambda \omega_j}{\|\omega\|}$$

$$\Rightarrow \nabla \lambda \|\omega\| = \frac{\lambda \vec{\omega}}{\|\omega\|}$$

$$\left(\sum_{i=1}^m (\langle \omega, x_i \rangle - y_i)^2 \right)'_{\omega_j} = \sum_{i=1}^m \left((\langle \omega, x_i \rangle - y_i)^2 \right)'_{\omega_j} = \sum_{i=1}^m \left(\left(\sum_{\ell=1}^d \omega(\ell) x_i(\ell) - y_i \right)^2 \right)'_{\omega_j}$$

$$= \sum_{i=1}^m 2 \left(\sum_{\ell=1}^d \omega(\ell) x_i(\ell) - y_i \right) \cdot x_i(j) = \cancel{2x_i(j)}$$

$$= \sum_{i=1}^m 2x_i(j) (\langle \omega, x_i \rangle - y_i)$$

$$\Rightarrow \nabla \left(\sum_{i=1}^m (\langle \omega, x_i \rangle - y_i)^2 \right) = 2 \sum_{i=1}^m x_i (\langle \omega, x_i \rangle - y_i)$$

$$\Rightarrow \boxed{\omega^{(t+1)} = \omega^{(t)} - \eta \left(\frac{\lambda \omega^{(t)}}{\|\omega^{(t)}\|} + 2 \sum_{i=1}^m x_i (\langle \omega^{(t)}, x_i \rangle - y_i) \right)}$$

(3) ב) -100 שנים. כפי שהי, $\nabla R(\omega^{(t)})$, SGD של η ו- λ

$$\nabla R(\omega^{(t)}, (x_i, y_i)) = \nabla (\langle \omega, x_i \rangle - y_i)^2 = 2x_i (\langle \omega, x_i \rangle - y_i)$$

$$\Rightarrow \boxed{\omega^{(t+1)} = \omega^{(t)} - \eta \left(\frac{\lambda \omega^{(t)}}{\|\omega^{(t)}\|} + 2x_i (\langle \omega^{(t)}, x_i \rangle - y_i) \right)}$$

(זאת כאן פה $\{x_1, \dots, x_m\}$ ו- y_i)

$\mathcal{H}_n = \Sigma_{i=1}^n \mathcal{B}_i^{-1}$ $X = \{0, 2\}^d$ 17 (4)

• n rows of p columns "x(i)=1" down entire PT

$$|\chi_n| \leq (d+2)^{(2^{n+1}-1)} \Rightarrow \text{rk}_n(a)$$

ב-13 במרץ 1964, פרסם המגזין *Time*, כי "האנדרטת הברזל" היא פרי יצירתו של...

באור יסוד לא תימר (2d) אכסר"י לא צוהר ב"ר.

– לכול ביטוי, קטעון, אל דימז נ יס אל קי-1 2^{n+1} פ.מ.3

– ניתן לראות בזה שהמחיר של המוצר הוא 10 שקלים.

לפי חוקי המבחן $n > 1$ כל המספרים יהיו זרים.

(על פי פירוש חזקוני) $|H_n| \leq (d+2)^{2^{n-1}}$ וכן הלאה.

(ב) יהי ERM וקור H_1 מדידת מידת הפער בין ERM ל- ERM . (נימוק: $d=3$)

$\text{ind}_{h_1, h_2}^{\text{acc}}(h, D) = 0.1$. יהי \hat{h} היחידה המינימלית כזו.

ראש נותן כי מחיר בקנה האטום כיון - $\inf_{h \in H} \text{cost}(h, D) = 0.1 + 0$

אם בעולם הזה? אגנוסטיק PAC learning

: Pⁿ דה מ פול D גורמא פל. H' = H₀, J = 0.05 E = 0.2 נוס

$$\frac{2(\log(1/4) + \log(2/5))}{\epsilon^2} = \frac{2(\log(5^{-3}) + \log(2/0.05))}{(0.2)^2} = 291 \leq m$$

התקום של האירוע ERM של שני נקודות בקו m $S \sim D^m$

הקאזורים יניק כסאטא קאק"א:

$$\text{err}(\hat{h}_3, D) \leq \inf_{h \in \mathcal{H}_0} \text{err}(h, D) + \varepsilon = 0.1 + 0.2 = 0.3$$

$$1 - \alpha = 1 - 0.05 = 95\%$$

הקטגוריה של אבות

$$\delta = 0.05, \varepsilon = 0.2 \in (0, 1)$$

נבחרין כי יגדלו קהילה אכן למקדמים:

$$|\mathcal{H}_y| \leq (d+2)^{\binom{2n+1}{2}+1} = 5^{33}$$

[illegible]