

8/6/20 - ↗

10 $\int_{\mathbb{R}^n}$

Model Selection : 2) 103) 52202

Feature Selection

יער אחד, עם אחד מיניהם נזכר - לא רק במונחים פיזיולוגיים אלא גם ב

הנחתה מינימלית גוררת עלייה ב- $f(x) - y$ כיוון שורדר

(λ גודל יחסית) Bias-complexity tradeoff - רעיון זה מגדיר

• תירטש, מ"מ תיירטש

ב- β_{1N} מ- α_3 ו- β_3 , ב-1: ρ_1 נ- ρ_2 ו- ρ_3

ר. פונוליסטי $\left\{ \begin{array}{l} \text{Train-Validation-Test - סגנון בז' (ii)} \\ \text{k-fold Cross Validation (iii) \rightarrow (训, בז' נעלם (ii)-בז')} \end{array} \right.$
 ר. אלטניטיבי $\left\{ \begin{array}{l} \text{ס. בז' נעלם מבחן ב- k-fold (iii)} \\ \text{ה- k-fold ב- k-fold (iv)} \end{array} \right.$
 : ר. בז' נעלם (2)

filter methods (i)

Forward & Backward - Stepwise Selection (ii)

Feature Transformations (iii)

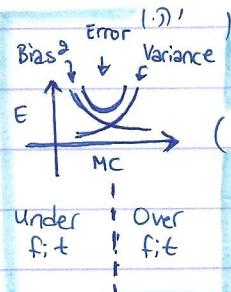
Model Selection (1)

የዚህ የሚከተሉት ስምዎችን እና በመስጠት ንብረቱ ማረጋገጫ ነው፡፡

$\mathbb{S} \in \mathcal{P}^N$ if \mathbb{S} is ERM_{H_i}(\mathbb{S}): (most likely) \mathbb{S} is an ERM

- Validation set = 20% of the data (FRM = 20% of the data)

$$h^* \in \text{ERM}_{\mathcal{H}}(V) \quad | \mathcal{B}_{N1} \quad \text{iid} \quad V = \{(x_i, y_i)\}_{i=1}^{m_V} \quad P(\cdot)$$



הנחות ורעלים

KNN -> K nearest neighbors 1. א.ר:

Soft-SVM -> λ regularization 2.

אחוז של כוונון 3.

אחוז של מילוי 4.

CART (Classification and Regression Tree) 5.

(רכזת החלטה)

(בגדיר פונקציית און) bagging - נńska און, T. 1: Bagging

(הבריגטס המשולבים) רפליקט ברג'סיה (כפנית און, k. 2

Adaboost (ריבוד און, T. 1: Boosting)

Best Subset Selection (בחירה מושלמת און, λ. 1: בחירה מושלמת)

Ridge (רכזת החלטה כעומק כפנית און, λ. 2

Lasso (רכזת החלטה כעומק כפנית און, λ. 3

: מושג לאם נורט ←

Train-Validation-Test Split (i)

הפרדה בין טעינה ותבנית: 'ה 3-ט טעינה ותבנית'

! מילוי

: מושג לאם נורט כפנית כפנית און, λ. 1

• $\hat{h}_\theta = \text{argmin}_{\theta \in \Theta} J_{\text{train}}(\theta)$

$h_\theta = A(S, \theta)$ פונקציית

• $J_{\text{train}}(h_\theta)$ פונקציית פיטר

$h^* = h_{\theta^*}$ s.t. $\theta^* = \text{argmin}_{\theta \in \Theta} J_{\text{train}}(h_\theta)$ פונקציית

לפניהם נורט כפנית כפנית און, λ. 2: מילוי

• $J_{\text{train}}(h^*)$ פונקציית פיטר

• $I_T(h^*)$ פונקציית פיטר

רתקן (NN) D מושג לאם נורט כפנית כפנית און, λ. 3

• פונקציית פיטר (NN) D מושג לאם נורט כפנית כפנית און, λ. 4

? מילוי?

המקרה הראשון - (ה 3-ט טעינה ותבנית און, λ. 1: מילוי)

train set -> מילוי (טבון) → overfitting → פיטר

• פונקציית פיטר (טבון) → λ. 2: מילוי

רְגִזָּה : (וְלֹא תַּנְאַתֵּן לְבָנֶךָ) תְּמֻנָּה לְבָנֶךָ וְלְבָנֶךָ
 lf מִתְּמֻנָּה כִּי כָּל הַמִּתְּמֻנָּה שֶׁל בָּנָם כָּל הַמִּתְּמֻנָּה
 lf מִתְּמֻנָּה כִּי כָּל הַמִּתְּמֻנָּה שֶׁל בָּנָם כָּל הַמִּתְּמֻנָּה

? בזבז י'pp 3-5 פון גולץ מכתבו ל'ג' גראן גראן מילר.

K-folds Cross Validation

• ANONYMOUS AUTHORSHIP AND THE PRACTICE OF PIRACY IN 18TH-CENTURY ENGLAND

ک اون * س پنلک پیزئن *: گپ

תְּבִזְבֵּחַ שָׂעִיר לְעֹלָה תְּמִימָה כְּלָמָד הַנְּצָרָת * A 3 נִסְטָח *

s_1, \dots, s_k ($s_1 = \dots = s_k$) op k-1 s $\sqcup_{1 \leq i \leq k}$ $p_i \geq 3N$ \wedge $\exists j$. 1
 $\vdash_{i \in [k]} B_i \bullet \vdash_{\theta \in \Theta} B_\theta$ 2

$$\text{Error}_i = L_{\hat{s}_i}(h_{i,\theta}) \quad \text{pln}_j \quad (\hat{s}_i \text{ fügt } A \text{ zu } \text{pln}_j)$$

$$\text{Error}(\theta) = \frac{1}{K} \sum_{i=1}^K \text{Error}_i : \text{Nur } \text{für } \text{pln}$$

$$\theta^* = \underset{\theta \in \Theta}{\operatorname{arg\,min}} \text{Error}(\theta) \quad \text{Nur } \text{für } \theta$$

: k second (verb)

CV \propto $\sigma_{\text{ref}} N^{\frac{1}{2}}$ $\because k=1$

PL 6P1 g-f nGIC3D 1/c pPSDN "Split-Sample CV" : K=2

. ስ.ያዥን ቅጂ የግብርና አድለነት ቅጂ የግብርና

("leave one out cv" 10 pu) k=m 1pl n3p7 n1pN2

רשות רגולציה וניהול

31PF 7301 Kd .10 = K 11C S = K rf 31PF 601P N - K 11D N

לכדרי יוניברסיטט (רשות הון) כוונת קאנדייר (הסנאט) נספחה ב-

• **PINNED** (knew) if I saw the number and it was pinned up.

Validation methods selected

iii

$$I_D(h_s) = \underbrace{I_D(h_s)}_A - \underbrace{I_V(h_s)}_B + \underbrace{I_V(h_s)}_B - \underbrace{I_S(h_s)}_S + \underbrace{I_S(h_s)}_E$$

. 11.33. (11) If loss \rightarrow 1% D overnight If loss \rightarrow 1% then - A

וְיַעֲשֵׂה כָּל-מַעֲשֶׂיךָ כִּי-כֵן תֹּהֵן

• INTK If loss \rightarrow IPB a,33fim If loss \rightarrow IP e,22n, - B

| N(k) f_b loss \rightarrow - c

Underfitting \Leftarrow פ. פ. פ. כ. - ; נ. פ. כ. י. ה. ב. - ; פ. א. נ. כ. פ. י. ו. כ. ו.

נירט נירט ו sk, ו lfpig lfjlc, lf pig 31ku 131nd plc

Overfitting ↙ גיינטיזציה

★ כוּלְכָמִים בְּבֵין כָּלֶבֶת וְבַשְׂרָבֶן כְּפָנֵי קָרְבָּן

accuracy

iv

$$\mathbb{E}[I_p(h_s)] = \frac{1}{m_v} \sum_{i=1}^{m_v} \mathbb{E}[I_p(h_s(x_i, y_i))] = I_p(h_s)$$

$$\mathbb{E}_{\substack{v \sim D^{mv}}} [\mathcal{I}_v(h_s)] = \frac{1}{m_v} \sum_{i=1}^{m_v} \mathbb{E}_{\substack{v \sim D^{mv}}} [l(h_s, (x_i, y_i))] \xrightarrow{\text{(DNC3), iid}} v$$

$$= \gamma_{m_v} \sum_{i=1}^{m_v} \mathbb{E}_{(x_i, y_i) \sim D} [L(h_s, (x_i, y_i))] = J_D(h_s)$$

• גַּתְּהֵן הַמִּזְרָחָה כְּבָנָה נֶגֶד אֶלְעָמֵן

: (ענברון עילית) רג'יסטר הרכבת NC plans וווער

$$\mathbb{P} \left[|I_\gamma(h) - I_p(h)| \geq \varepsilon \right] \leq g \cdot e^{-2M_\gamma \varepsilon^2}$$

$\underbrace{}_{\mathbb{E}[I_\gamma(h)]} \quad : \varepsilon$

$$\therefore \epsilon = \sqrt{\frac{1}{2M_V} \cdot \ln\left(\frac{g}{\delta}\right)}$$

$$\mathbb{P}[|\mathcal{I}_V(h) - \mathcal{I}_p(h)| \leq \sqrt{\frac{\log(\frac{2}{\delta})}{g m_V}}] \geq 1 - \delta$$

(p_{1.3} - \varepsilon, \text{confidence } \delta, \text{ all } \omega \in \Omega \text{ on } M_V)

קווין בז'רנו (2)

(P.1) נסמן \mathcal{S} כSubset של \mathcal{X} ו- \mathcal{B}_S כSubset של \mathcal{B} .
 מינימיזציה של פונקציית האפסון מושגת על ידי $\mathcal{S} = \mathcal{B}_S$.

המשמעות של \mathcal{S} היא \mathcal{S} הוא subset המינימיזה פונקציית האפסון. \mathcal{B}_S הוא best subset (best subset \rightarrow subset שמיינימיזה פונקציית האפסון).
 מינימיזה פונקציית האפסון מושגת על ידי $\mathcal{S} = \mathcal{B}_S$.

ריצוף - filter Methods (i)

כדי למדוד את הדרישות הדרישה לsubset \mathcal{S} מושגת באמצעות scoring function.

$$\text{score}(\mathcal{S}) = \min_{a,b \in \mathbb{R}} \frac{1}{m} \sum_{i=1}^m (a_i(x_i)_i + b - y_i)^2$$

a bias \rightarrow subset \mathcal{S} מושגת באמצעות a ו- b מינימיזה פונקציית האפסון. מינימיזה פונקציית האפסון מושגת על ידי $\mathcal{S} = \mathcal{B}_S$.

$$\text{score}(\mathcal{S}) = \min_{a,b \in \mathbb{R}} \frac{1}{m} \| a \cdot v^{\mathcal{S}} + b \cdot \mathbf{1}_m - y \|_2^2$$

y ו- $v^{\mathcal{S}}$ הם m מטרים כיוון ש- b הוא bias \rightarrow $b = j$ ו- $v^{\mathcal{S}} = v_j$.

לפיכך $\text{score}(\mathcal{S}) = \text{score}(v_j)$

לפיכך $\text{score}(\mathcal{S}) = 1$

$$\text{score}(\mathcal{S}) = \min_{a,b \in \mathbb{R}} \frac{1}{m} \| a \cdot v^{\mathcal{S}} - b \cdot \mathbf{1}_m + b \cdot \mathbf{1}_m - (y - \bar{y} \cdot \mathbf{1}_m) \|_2^2$$

$$v^{\mathcal{S}} = \frac{1}{m} \sum_{i=1}^m (v^{\mathcal{S}})_i, \quad \bar{y} = \mathbf{1}_m \cdot \frac{1}{m} \sum_{i=1}^m y_i$$

מינימיזה פונקציית האפסון מושגת על ידי $\mathcal{S} = \mathcal{B}_S$.
 מינימיזה פונקציית האפסון מושגת על ידי $\mathcal{S} = \mathcal{B}_S$.

$$(I) \frac{\partial}{\partial b} \|a \cdot (v^j - \bar{v}^j \cdot 1m) + b \cdot 1m - (y - \bar{y} \cdot 1m)\|_2^2 = 2b m \quad : \text{if } k=j : \text{gekennzeichnet}$$

$$(II) \frac{\partial}{\partial a} \|a \cdot (v^j - \bar{v}^j \cdot 1m) + b \cdot 1m - (y - \bar{y} \cdot 1m)\|_2^2 = : \text{seit } b=0 \text{ platzieren}$$

$$= 2a \|v^j - \bar{v}^j \cdot 1m\|^2 - 2 \langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle$$

$b=0$ - es folgt a ist ein Koeffizient von v^j

: (PP) obiges gilt für $j \neq k$ ($b=0$ weiter) a ist ein Koeffizient von v^j

$$\therefore a = \frac{\langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle}{\|v^j - \bar{v}^j \cdot 1m\|_2^2}$$

$$\langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle$$

$$a = \frac{\|v^j - \bar{v}^j \cdot 1m\|_2^2}{\|x - y\|_2^2}, b = 0 \quad (\text{SINNN}) \quad \text{gesucht } y \text{ mit } \|x - y\|_2^2 = \|x\|_2^2 - 2 \langle x, y \rangle + \|y\|_2^2$$

. (oder $b \rightarrow y = y - \bar{y} \cdot 1m, x = a(v^j - \bar{v}^j \cdot 1m)$)

$$\|x - y\|_2^2 = \langle x - y, x - y \rangle = \langle x, x \rangle - 2 \langle x, y \rangle + \langle y, y \rangle = : \text{normiert werden}$$

$$= \|x\|_2^2 - 2 \langle x, y \rangle + \|y\|_2^2$$

$$\text{score}(j) = \underbrace{\frac{1}{m} \left(\|y - \bar{y} \cdot 1m\|_2^2 - \langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle^2 \right)}_{\leftarrow j-\text{Punkt-Koeffizient}} : \text{(PP)}$$

$$\text{score}(j) = \frac{-\langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle}{\|v^j - \bar{v}^j \cdot 1m\|_2^2} : \|y - \bar{y} \cdot 1m\|_2^2 \cdot -P \text{ prüft, wenn nicht } \frac{1}{m} \text{ normiert werden}$$

$$\text{score}(j) = \frac{-\langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle}{\|v^j - \bar{v}^j \cdot 1m\|_2^2 \cdot \|y - \bar{y} \cdot 1m\|_2^2}$$

gilt für $j \neq k$ $\text{score}(j)$ ist gleich $\text{score}(k)$ \Rightarrow $\text{score}(j) = \text{score}(k)$ für alle $j, k \in \{1, 2, \dots, m\}$

$$\langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle$$

$$\text{score}_3(v^j) = \left| \frac{\langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle}{\|v^j - \bar{v}^j \cdot 1m\|_2^2 \cdot \|y - \bar{y} \cdot 1m\|_2^2} \right| =$$

$$= \left| \frac{\frac{1}{m} \langle v^j - \bar{v}^j \cdot 1m, y - \bar{y} \cdot 1m \rangle}{\sqrt{\frac{1}{m} \|v^j - \bar{v}^j \cdot 1m\|_2^2} \sqrt{\frac{1}{m} \|y - \bar{y} \cdot 1m\|_2^2}} \right| = \dots =$$

$$= \left| \frac{\frac{1}{m} \sum_{i=1}^m ((x_i)_j - \bar{v}^j) (y_i - \bar{y})}{\sqrt{\frac{1}{m} \sum_{i=1}^m ((x_i)_j - \bar{v}^j)^2} \cdot \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2}} \right|$$

אנו מגדירים קוריאנס כפונקציה של סכום הערך הממוצע מ- x_i ו- y_i (בנוסף ל- σ^2):

Pearson Correlation coefficient r_{ij} הession
 $i \text{ first}, j \text{ second}$

318) רעלס גוּאַלְעָדֶן אַפְּרִיאַלְעָדֶן, קָרְבָּנְיָהָן, וְקָרְבָּנְיָהָן

- (11) מי יתיר על מילוי תפקידו? ?
 מי יתיר על מילוי תפקידו?

רְכָב כַּמְרֵב הַיּוֹצֵא גַּדְעִין אֶלְגָּיָה :

Forward & Backward Stepwise Selection

3nlc-3nlc גדרוני רג'סט קס גדרון - Forward : 11:130 ור'.

• பி.ஃ.ா குஜரை 36, நல்ல பிள்ளை 13.00 ரூ.

3nf-3nf ניגודן פיזיון ב יי ה.פ.ן - Backward

የኢትዮጵያውያንድ አስተዳደር የሚከተሉ ስራው ተስፋል

: best subset select. 亂子 亂. NPG. 亂 亂 亂

1. פטישים וריגושים נאורים וארוכים.

. β_{INT} מינימיזירט את הערך של $\hat{\sigma}^2$ על ידי best subset . &

• **الخطابات** هي عبارة عن رسائل مكتوبة توجه إلى الآخرين، وتهدف إلى نقل معلومات أو إرادة أو تشريع.

best subset- δ $P_{\text{avg}}/P_{\text{PP}}$ P_{N13} P_{D} η_{defNS} *

: ၇၃၅၁၇၁၂၄

הציגו יסודות נספחים למסמך.

רְבָרִין, מיל' צוֹנֶה לְהַזְמִין כְּנָסֶר וְלִבְנָה, אֲבִיב

בגדיים דומים נסבטיים (בנוסף לארון) מוצאים רישום של א-ספתקים וריבויים.

: מנגנון אובייקטיבי

Feature Transformation iv

רואה את ה- f_i , ובלצטוטה, מנסה לשים אותו בטווח מסוים.

: מנגנון אובייקטיבי

לעומת זה, מנגנון אובייקטיבי מנסה לבודד את המבנה של f_i .

$$f_i \leftarrow f_i - \bar{f} : \text{מנגנון אובייקטיבי}$$

$$\left(\bar{f} = \frac{1}{m} \sum_{i=1}^m f_i, i \rightarrow \text{ה-}i^{\text{th}} \text{ מרכיבי אוניברסיטאי} \right) \text{ מנגנון אובייקטיבי}$$

לעומת זה, מנגנון אובייקטיבי מנסה לבודד את המבנה של f_i .

$$f_{\min} = \min_j f_j, f_{\max} = \max_j f_j \quad \text{מן}: [0, 1] \quad \text{לפניהם}: \begin{matrix} 1 \\ 3 \\ 2 \\ 6 \end{matrix}$$

$$f_i \leftarrow \frac{f_i - f_{\min}}{f_{\max} - f_{\min}} : \begin{matrix} 1 \\ 3 \\ 2 \\ 6 \end{matrix}$$

לעומת [-1, 1] מנגנון אובייקטיבי מנסה לבודד את המבנה של f_i .

$$f_i \leftarrow 2 \frac{f_i - f_{\min}}{f_{\max} - f_{\min}} - 1 : \text{מנגנון אובייקטיבי}$$

לעומת [-1, 1], מנגנון אובייקטיבי מנסה לבודד את המבנה של f_i .

$$f_i \leftarrow \frac{f_i - \bar{f}}{\sqrt{\frac{1}{m} \sum_{i=1}^m (f_i - \bar{f})^2}} : \text{מנגנון אובייקטיבי}$$

: מנגנון אובייקטיבי