

מבוא לבינה מלאכותית – תרגיל 1:

1. (על תרגול 1) נתונות התצפיות הבאות, המחולקות ל-training ול-test:

training:

[(12.47, 1.12), (6.565, 0.795), (0.73, -0.004), (5.608, 0.802), (9.75, -0.995), (6.668, 0.81),
(12.271, 0.986), (14.338, -0.527), (5.72, 1.013), (8.805, -0.759), (16.204, 0.459),
(7.219, 0.476)]

test:

[(3.974, 0.83), (0.114, -0.145), (7.001, 0.703), (9.894, -0.897), (7.053, 0.756), (4.243, 0.765),
(13.82, -0.38), (4.484, 0.998)]

- א. השתמשו בפונקציית loss ריבועית ללא רגולריזציה.
בצעו רגרסיה פולינומיאלית בהתבסס על הנקודות מה-training בלבד לפולינומים מהמעלות הבאות: 2, 6, 16.
- מהם ערכי ה- γ המתאימים לנקודות מה-test? מהו ה-loss?
- ב. כעת, השתמשו בפונקציית loss עם רגולריזציה L2 (ridge), עבור רגרסיה לפולינום ממעלה 6 עם מקדם רגולריזציה λ . כתבו פונקציה המחשבת את ה-loss על ה-training ועל ה-test כתלות במקדם λ וציירו אותה.

התרגיל הזה נועד לשתי מטרות. האחת: להכניס טוב יותר לראש את המושגים מהתרגול הראשון. השנייה: זה תרגיל תכנותי לא מסובך במיוחד ב-python, לכן למי שהשפה חדשה לו זה יהיה נחמד. מוזמנים להיעזר [בקישור הבא](#).

באופן כללי, אני ממליץ להכיר את pycharm כסביבת עבודה לתכנות ב-python. בנוסף, ספריות חשובות (לפחות לשלב זה) הן numpy ו-scikit-learn.

2. (על תרגול 2)

א. חשבו את מימד VC של $\mathcal{H} = \{h_{a,b}(x) = \mathbb{1}_{(a,b)}(x) \mid a < b, a, b \in \mathbb{R}\}$ על \mathbb{R} .

ב. הראו שמימד VC של מחלקת השערות סופית \mathcal{H} (לתיוג בינארי) חסום ע"י $\log_2 |\mathcal{H}|$.

ג. הראו שמימד VC של המחלקה הבאה

$$\mathcal{H} = \{h_{(w,b)}(x) = \begin{cases} 1, & w_1x_1 + w_2x_2 + b > 0 \\ -1, & \text{else} \end{cases} \mid (w, b) \in \mathbb{R}^3\}$$

על מרחב קלט \mathbb{R}^2 , הוא 3.

3. (על תרגול 2)

נתבונן בבעיה הבאה:

מרחב הקלט: $\{(0,0), (1,0), (0,1), (1,1)\}$

תיוגים מתאימים: $(0,0): 0, (1,0): 1, (0,1): 1, (1,1): 0$ [פונקציית XOR].

מחלקת ההשערות: מסווגים לינאריים (אם $w \cdot x + b \geq 0$ אז החיזוי יהיה 1 ואחרת 0).

האם ניתן להגיע לסיווג מושלם? אם כן, מצאו את הפרמטרים לכך. אם לא, נסו לחשוב על דרכים איך להגיע לכך שיהיה לבעיה פתרון.