

האם יש צורך בלמידה מונחית?

האם יש צורך בלמידה מונחית?

$$\text{Gain}(S, A) = H(S) - \sum_{v \in \text{value}(A)} \left(\frac{|S_v|}{|S|} \cdot H(S_v) \right) \quad \text{Information Gain}$$

האם יש צורך בלמידה מונחית?

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	hot	high	Weak	No
2	Sunny	hot	high	Strong	No
3	Overcast	hot	high	Weak	Yes
4	Rain	mild	high	Weak	Yes
5	Rain	Cool	normal	Weak	Yes
6	Rain	Cool	normal	Strong	No
7	Overcast	Cool	normal	Strong	Yes
8	Sunny	mild	high	Weak	No
9	Sunny	Cool	normal	Weak	Yes
10	Rain	mild	normal	Weak	Yes
11	Sunny	mild	normal	Strong	Yes
12	Overcast	mild	high	Strong	Yes
13	Overcast	hot	normal	Weak	Yes
14	Rain	mild	high	Strong	No

האם יש צורך בלמידה מונחית?

האם יש צורך בלמידה מונחית?

האם יש צורך בלמידה מונחית?

האם יש צורך בלמידה מונחית?

האם יש צורך בלמידה מונחית?

המקרה שבו יש דבר אחד

$$H(\text{Play-Tennis}) = \sum P_i \log P_i$$

$$= 0.84$$

הערך הממוצע של המידע הזה הוא 0.84

לפי התוצאה הממוצעת הזו יש לנו סתם 1.6 ביט
כמה מידע יש לנו? קודם כל יש 9 מידע
ולכן 5 ביט - 1.6 = 3.4 ביט

$$S: [8^+, 5^-]$$

סדרה של 8 ו-5

אם כן, הרי שיש לנו 1.6 ביט - 0.84 = 0.76 ביט
נראה שיש קשר בין המידע הזה לבין המידע
האחר. יש לנו 1.6 ביט - 0.84 = 0.76 ביט
אם כן, נראה שיש קשר בין המידע הזה לבין המידע
האחר.

$$S_1: [3^+, 4^-]$$

הערך הזה

הערך הזה

אם כן, נראה שיש קשר בין המידע הזה לבין המידע

$$S_2: [6^+, 1^-]$$

אם כן, נראה שיש קשר בין המידע הזה לבין המידע

$$\text{Gain}(S, \text{Humidity}) =$$

play Tennis



$$= H(s) - \sum \frac{(\text{which-humidity})}{\text{prior prob}} \cdot H(\text{which humidity})$$

$$= H(s) - \left(\frac{(\text{humidity-high})}{151} \cdot H(\text{humidity-high}) + \frac{(\text{humidity-normal})}{151} \cdot H(\text{humidity-normal}) \right)$$

$$= 0.940 - \frac{7}{14} \cdot H(\text{humidity-high})$$

$$- \frac{7}{14} H(\text{humidity-normal})$$

per 15/0.96 per 151 14 per 151 14 per 151 14

$$H(\text{humidity-high}) = \left\{ \begin{array}{l} \text{per 15/0.96} \\ \text{per 151 14} \\ \text{per 151 14} \end{array} \right\}$$

$$= \left\{ \begin{array}{l} \text{per 151 14} \\ \text{per 151 14} \\ \text{per 151 14} \end{array} \right\} =$$

$$= 0.985$$

3+, 4- 1.01

per 151 14

$$H(\text{humidity-normal}) = 0.592$$

התוצאה היא 0.151

$$\text{Gain}(S, \text{humidity}) = 0.940 - \frac{7}{14} \times 0.985 - \frac{7}{14} \times 0.592 = 0.151$$

התוצאה היא 0.151
התוצאה היא 0.151
התוצאה היא 0.151

התוצאה היא 0.151
התוצאה היא 0.151
התוצאה היא 0.151

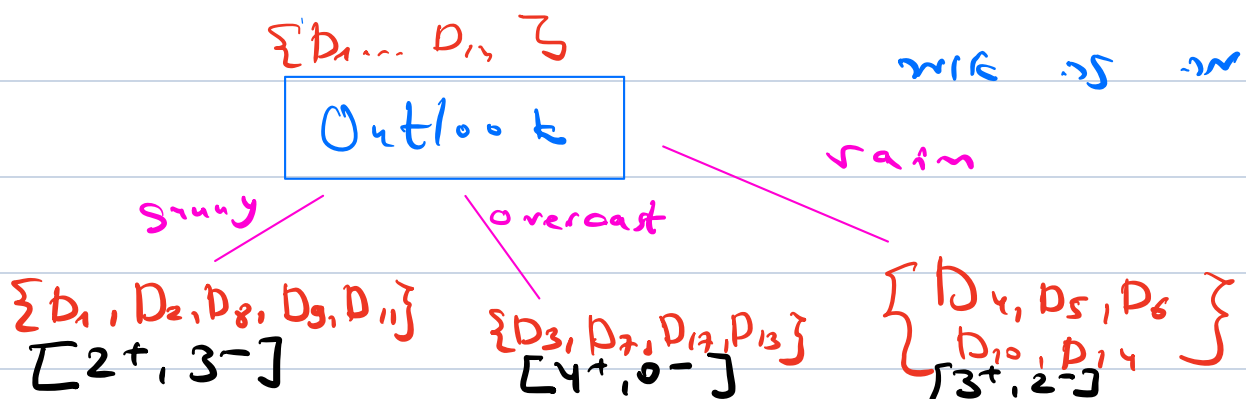
$$\text{Gain}(S, \text{wind}) = 0.94 - \frac{8}{14} \times 0.811 - \frac{6}{14} \times 1 = 0.48$$

התוצאה היא 0.48

$$\text{Gain}(S, \text{outlook}) = 0.241$$

$$\text{Gain}(S, \text{temperature}) = 0.029$$

התוצאה היא 0.241
התוצאה היא 0.241
התוצאה היא 0.241



התהליך של חישוב האינפורמציה
המקסימלית - (המקסימלית) היא 1 ביט

הערות: ישנם 16 מקרים.

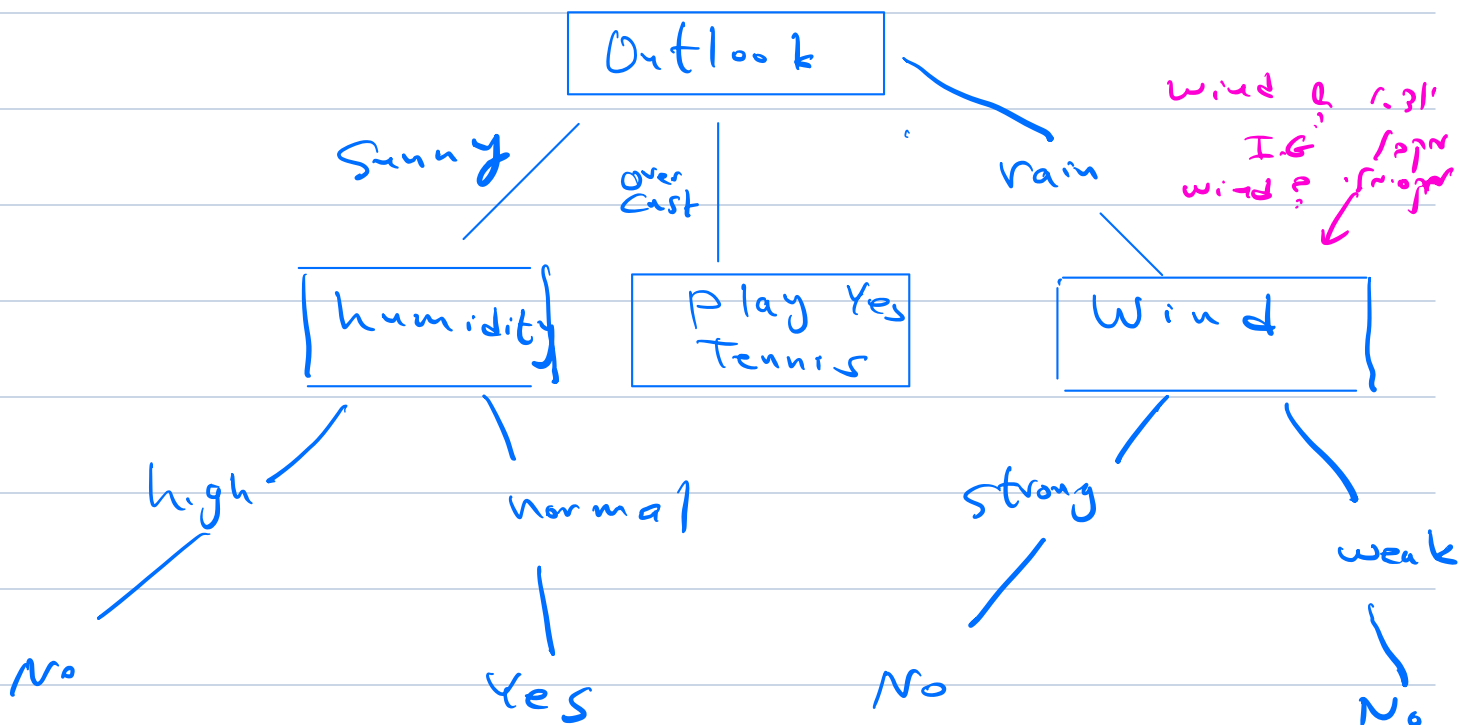
התוצאות של חישוב האינפורמציה
המקסימלית - (המקסימלית) היא 1 ביט
Tennis

$$\text{Gain}(\text{Sunny}, \text{humidity}) = 0.97 - \frac{3}{5} \log \frac{3}{5} - \frac{2}{5} \log \frac{2}{5}$$

$$\text{Gain}(\text{Sunny}, \text{Temperature}) = 0.570$$

$$\text{Gain}(\text{Sunny}, \text{Wind}) = 0.19$$

humidity? אינפורמציה - Gain - אחר
הוא 0.97



התוצאות של חישוב האינפורמציה

הצגת אלגוריתם:

היציאה בצד המאמץ היא לרוב אפילו כמעט
אשר כפי שהי אפילו היציאה לא סובלנית
על מודלים אחרים כפי שהי אפילו אפילו
אשר היציאה היציאה היציאה היציאה
אשר היציאה היציאה היציאה היציאה
אשר היציאה היציאה היציאה היציאה

כפי שהי אפילו היציאה היציאה היציאה
אשר היציאה היציאה היציאה היציאה

אלגוריתם Random Forest

אשר היציאה היציאה היציאה היציאה

אשר היציאה היציאה היציאה היציאה

אשר היציאה היציאה היציאה היציאה

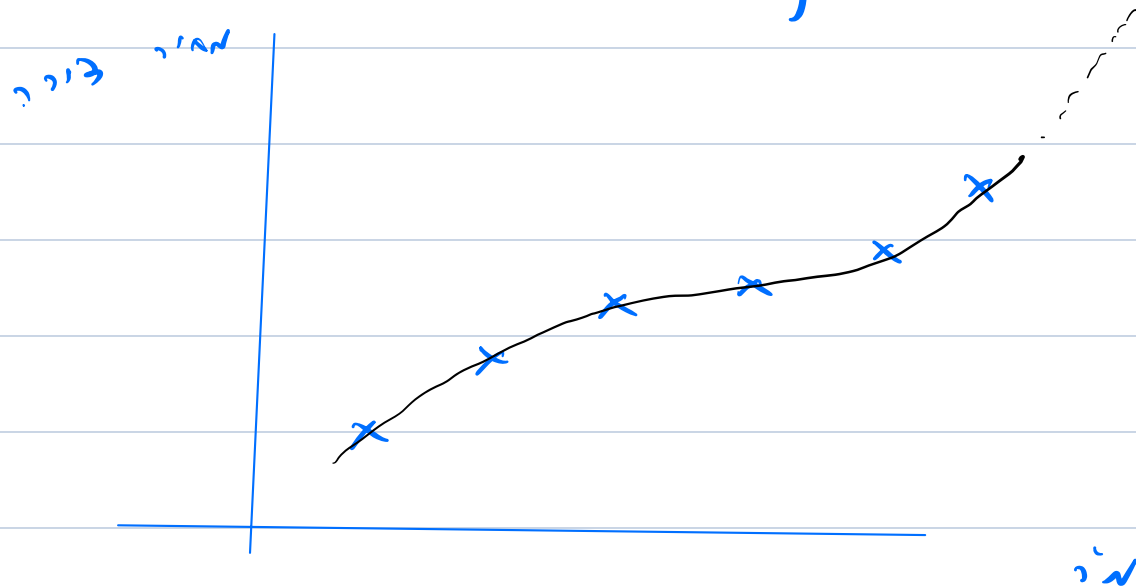
אשר היציאה היציאה היציאה היציאה

חלש קיבולו בלחץ של עצי האלון, וכדי לסלק
 צומח א נבחר מה העץ השטח של $\tilde{x}$ בל
 העצמי.

אם עכשיו ציגינו את סיווג קטגוריאלי, בדיוק על
 בנאלי, אן יש לנו כמה רציונל-אן אהלים העין
 זה אמר בחלק-?

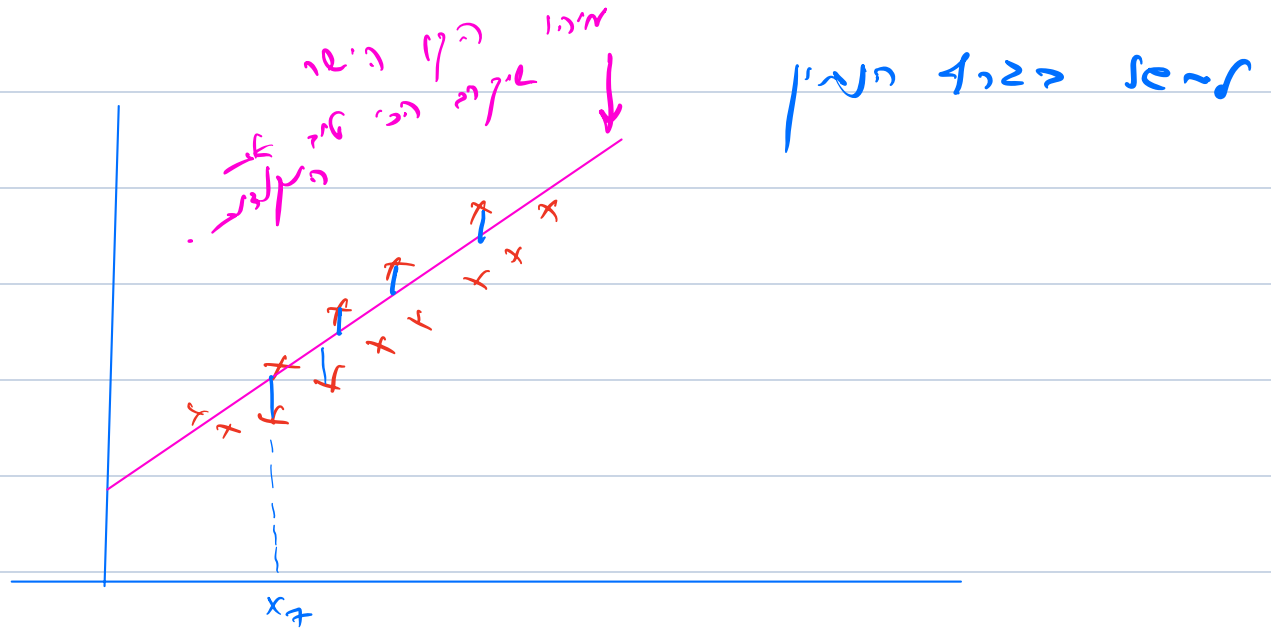
אילו אומר העין העין לעל- סיווג האשן נלך
 האש וס סיווג שני נלכ האש וכן הלאה.
 נבחר בהמשך הקוים של- נוסט.

אמר להילכו ש- $\tilde{x}$ קטגוריאלי-אליה העלה
 זה רשמי לעין? R^n . לצומח $\tilde{x}$ בל



במקום אחר ציגנו הקלטה נוסט לעל אש קיורק
 תצורה, בהם שטח יחיד פולנומאלי אש
 העלה מלכס סכר (כנס לעלה במעלה

הסיכוי של overfitting יגדל.



רציונליות

מחיר המכירה נקבע על ידי מספר קטן של משתנים, אך יש הרבה גורמים המשפיעים על המחיר. לדוגמה, המיקום, המבנה, המצב, וכו'. לכן, יש צורך במודל שיכול להעריך את המחיר בצורה טובה יותר.

$$\{ \vec{x}_i, y_i \}_{i=1}^n$$

אנו נרצה לבנות מודל שיכול להעריך את המחיר בצורה טובה יותר. קיימת דרך אחת לעשות זאת, וזוהי דרך המינימום הריבועי (Least Squares).

$$\hat{y} = f(\vec{w}, \vec{x}) = w_0 + w_1 x_1 + \dots + w_d x_d$$

כאשר $\vec{w}$ הוא וקטור המשקלים, ו- $\vec{x}$ הוא וקטור המאפיינים.

עצם המודל שלנו ינסה למצוא את הווקטור $\vec{w}$ הטוב ביותר, כלומר, הווקטור שימזער את הריבוע של שגיאת התחזית.

הפעלה הקטנה ביותר של מודל 4000 על התקופה.
 הניסוח מייחס העדפה לטוב קטנה על התקופה, במידה
 אפשרית - מידת הפחיתות

הבעיה למציאת המשתנה

אם הישג הוא איננו, אזי הוא מוזנח

$\hat{y} = ax + b$ בקרה של קבוצה פרויקטית

מקורו של w וכן $w_0 + w_1 x$ הוא המשתנה המצוי

נניח כי הערך המיועד של y_i הוא $\hat{y}_i$
 והערך בפועל הוא y_i (הערך המצוי)
 שגיאת y שגיאת

$$e_i = |y_i - \hat{y}_i|$$

הקו הירוק מראה את השגיאה של הקו הירוק
 שגיאת העדפה היא מידת

$$\sum e_i = \sum |y_i - \hat{y}_i| = \sum \sqrt{(y_i - \hat{y}_i)^2}$$

כדי למצוא את המשתנה המצוי
 המשתנה, נניח כי המשתנה המצוי
 (הערך המצוי)

$$\hat{y} = w_0 + w_1 x$$

$$J(\vec{w}) = \sum_{i=1}^n e_i = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$= \sum_{i=1}^n (y_i - (w_0 + w_1 x))^2$$

יש להצטרף מינימום. ר"ל:

w_0, w_1

בין מוצאים מינימום באופן אנליטי, נגזרת
אם יש המגרמ w_0, w_1 ישנה אופסה
הנורמל להסיק והמינוס הימליס חזונה.

$$\frac{\partial J}{\partial w_0} = 2(y_i - (w_0 + w_1 x)) = 0$$

$$\frac{\partial J}{\partial w_1} = 2w_1(y_i - (w_0 + w_1 x)) = 0$$

כדי שזה יהיה מינימום בלי נאליזם / צריך
לחפץ w_0, w_1 ! w_1 , ו/או w_0 להצטרף?
אם הנגזרת מסתירה חזונה.

איפה עולה הקצה?

הישי לאנחנו בואו מנגד צריך ציטוט סימול
משה נתינה לנו ו צוהאלל לזה צריך
לע w_1 הישי האב בואו , ולכן האלף
האנליטי בדיק כולל על כן השל.

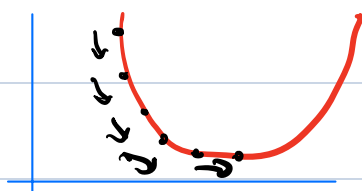
בשיטת ה- Gradient descent אנחנו חשבים את הפונקציה של θ ונמצא את המינימום שלה.

שיטת ה- Gradient descent

הינה שיטת חישוב פשוטה למציאת המינימום של פונקציה. θ - זהו המאפיין של הפונקציה. אנחנו מחשבים את הפונקציה ואת הנגזרת שלה. אנחנו ממשיכים עד שהנגזרת קרובה ל-0.

כדי למצוא את המינימום של פונקציה אנחנו מחשבים את הפונקציה ואת הנגזרת שלה. אנחנו ממשיכים עד שהנגזרת קרובה ל-0. אנחנו ממשיכים עד שהנגזרת קרובה ל-0. אנחנו ממשיכים עד שהנגזרת קרובה ל-0.

אנחנו ממשיכים עד שהנגזרת קרובה ל-0. אנחנו ממשיכים עד שהנגזרת קרובה ל-0. אנחנו ממשיכים עד שהנגזרת קרובה ל-0. אנחנו ממשיכים עד שהנגזרת קרובה ל-0.



כפי שזכרנו, חידוש קצוץ זה נקרא היום פז'יט
קצב ההתחממות הוא 4°C, כלומר, אטמוספירה
מ'צ', ציף אטמוספירה של 1°C, אטמוספירה
כמו שאמרנו מרחוק דאג'לה (1°C)

$\omega^{(0)} \rightarrow \omega^{(1)} \dots \omega^{(n)}$
 כן $\omega^{(n)} \rightarrow \omega^{(n+1)}$

בן זבדן סי מנג 00 א ד'ד'ל קל'ג'ג'ג'

$$\omega^{(t+1)} = \omega^{(t)} - \underbrace{\alpha}_{\text{Lernrate}} \cdot \frac{dJ}{d\omega} \bigg|_{\omega^{(t)}}$$

י' חנוכה אחר הצהריים

$\omega^{(k)}$ - ה' א' ו' / פלגים

הצגת המידע

הקבוצה α קרבה $\rightarrow$ צורת הצורה

α β γ δ ϵ ζ η θ ι κ λ μ ν ξ $\omicron$ π ρ σ τ υ ϕ χ ψ ω

[illegible]

$$\omega^{(t+1)} = \omega^{(t)} - \alpha \nabla J|_{\omega^{(t)}}$$

3437 3/4 — 2 1/2 2 1/2 3 1/2 3 1/2

ע' אונטערן געזעץ פון פארשטאנד.

3.06 זה Gradient descent זה משימה

$$f(x, y) = 4x^2 - 4xy + 2y^2 \quad \text{לפינו} \quad \text{לפינו}$$
$$w^{(0)} = (2, 3) \quad ? \quad \text{לפינו}$$

המשימה: זהו ————— המשימה —————
המשימה ————— המשימה —————
המשימה ————— המשימה —————
המשימה ————— המשימה —————

המשימה ————— המשימה —————
המשימה ————— המשימה —————
המשימה ————— המשימה —————

$$\nabla f(x, y) = (f'_x, f'_y) = (8x - 4y, -4x + 4y)$$

זהו המשימה

$$\nabla f(w^{(0)}) = \nabla f(2, 3) = (4, 4)$$

$$w^{(1)} = w^{(0)} - \alpha \nabla f|_{w^{(0)}} =$$

$$= (2, 3) - \alpha (4, 4)$$

$$= (2 - 4\alpha, 3 - 4\alpha)$$

זהו המשימה, המשימה

$$\alpha = 0.01$$

המשימה, המשימה

$$w^{(1)} = (2 - 4\alpha, 3 - 4\alpha)$$

$$w^{(1)} = (2, 3) - \alpha (4, 4)$$

אם $t = \alpha$ נמצא את המינימום של f

$$f(t) = f(2-4t, 3-4t)$$

אם f היא פונקציה של שני משתנים

$$f'(t) = -\nabla f(2-4t, 3-4t) \cdot (-4, 4)$$

$$= 64\alpha - 32$$

אם $f''(t) > 0$

אז f היא פונקציה קמורה

אם $f''(t) < 0$ אז f היא פונקציה קעורה

אם $f''(t) = 0$ אז f היא פונקציה קמורה או קעורה

$$w^{(2)} = w^{(1)} - \frac{1}{2} \nabla f(w^{(1)})$$

אם $w^{(1)} = (0, 0)$ אז $w^{(2)} = (0, 0)$

אם $w^{(1)} = (0, 0)$ אז $w^{(2)} = (0, 0)$