לא לשכוח להפעיל הקלטה!

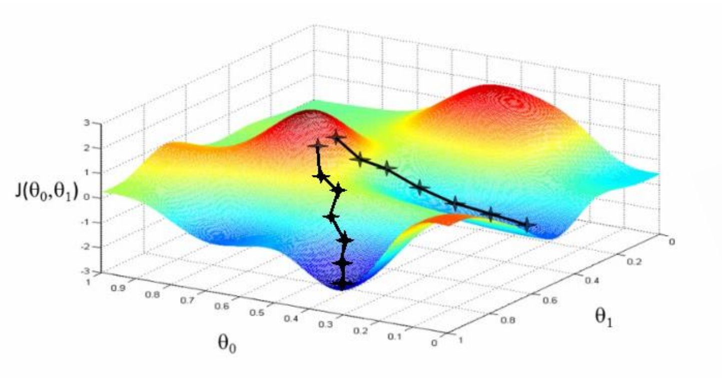
מה נלמד בשיעור זה?

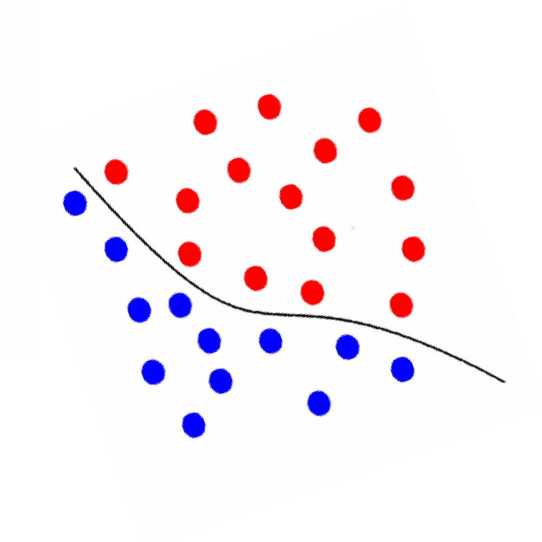
- ✓ פונקציה לוגיסטית
- ✓ Cost function לרגרסיה לוגיסטית
- ✓ gradient descent עבור רגרסיה לוגיסטית
- ✓ בעיות סיווג לשתי מחלקות

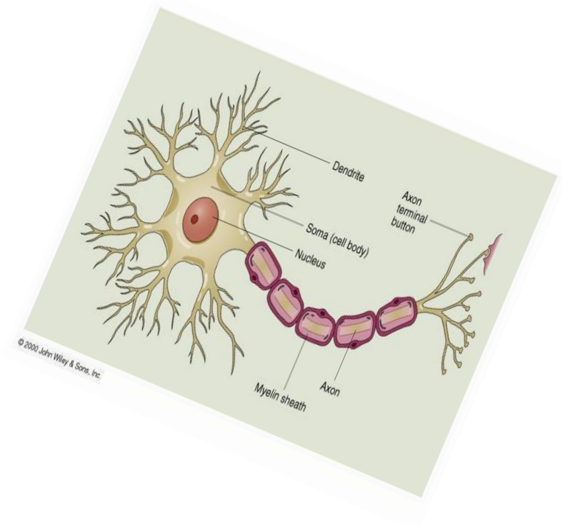
מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

מכון דוידסון לחינוך מדעי
DAVIDSON INSTITUTE OF SCIENCE EDUCATION

רגרסיה לוגיסטית







שיעור סינכרוני

רגרסיה לוגיסטית

- אלגוריתם סיווג
- מחשב את ההסתברות שדוגמא לא ידועה שייכת לכל אחד מהסוגים
- **שיטת החישוב:** מציאת ערכי המשקולות w ו-b כאשר ערך הטעות מפונקציית המחיר היא מינימלית

פונקציה לוגיסטית

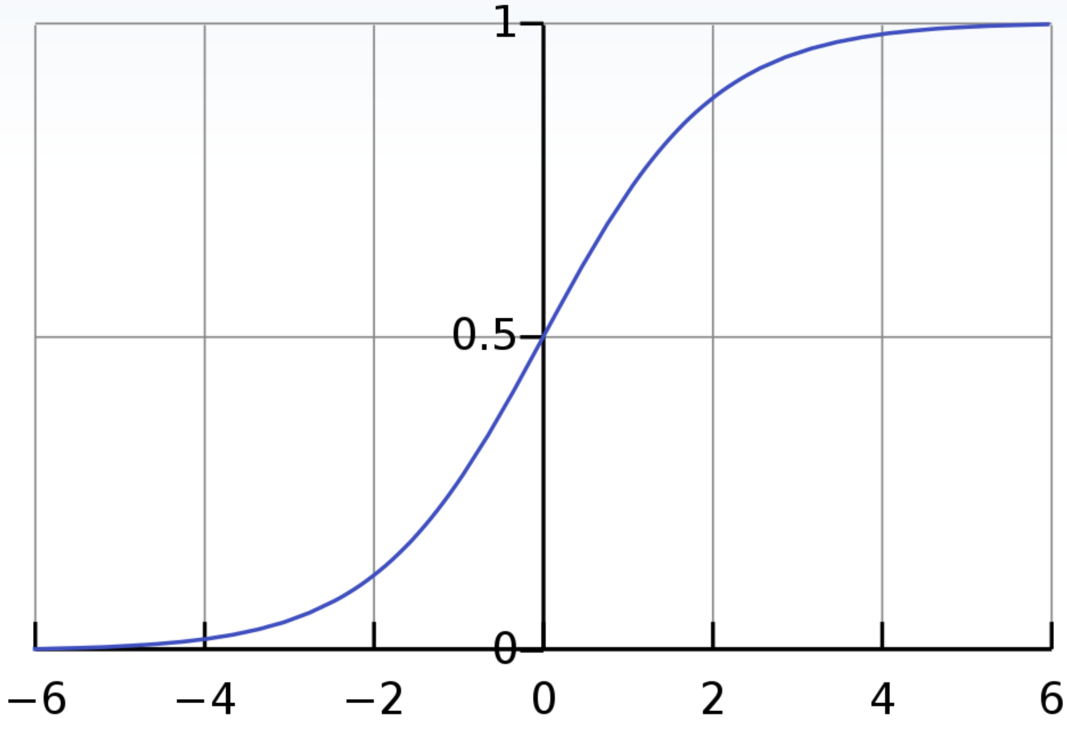
פונקציה לוגיסטית

$$\hat{y} = \frac{1}{1+e^{-z}}$$

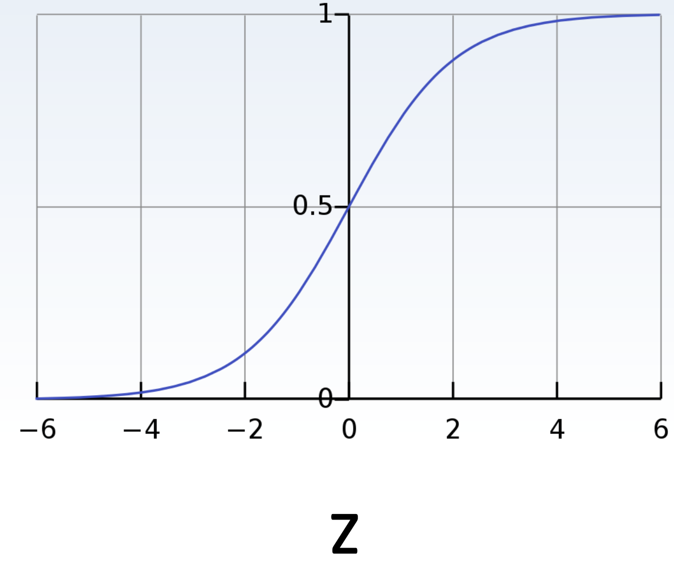
כאשר **z** הינו הקלט המשוקלל לנוירון

$\sum_{i=1}^{n} \mathbf{w}(\boldsymbol{i})\mathbf{x}(\mathbf{i}))$+b

n–מספר התכונות לא כולל הטיה

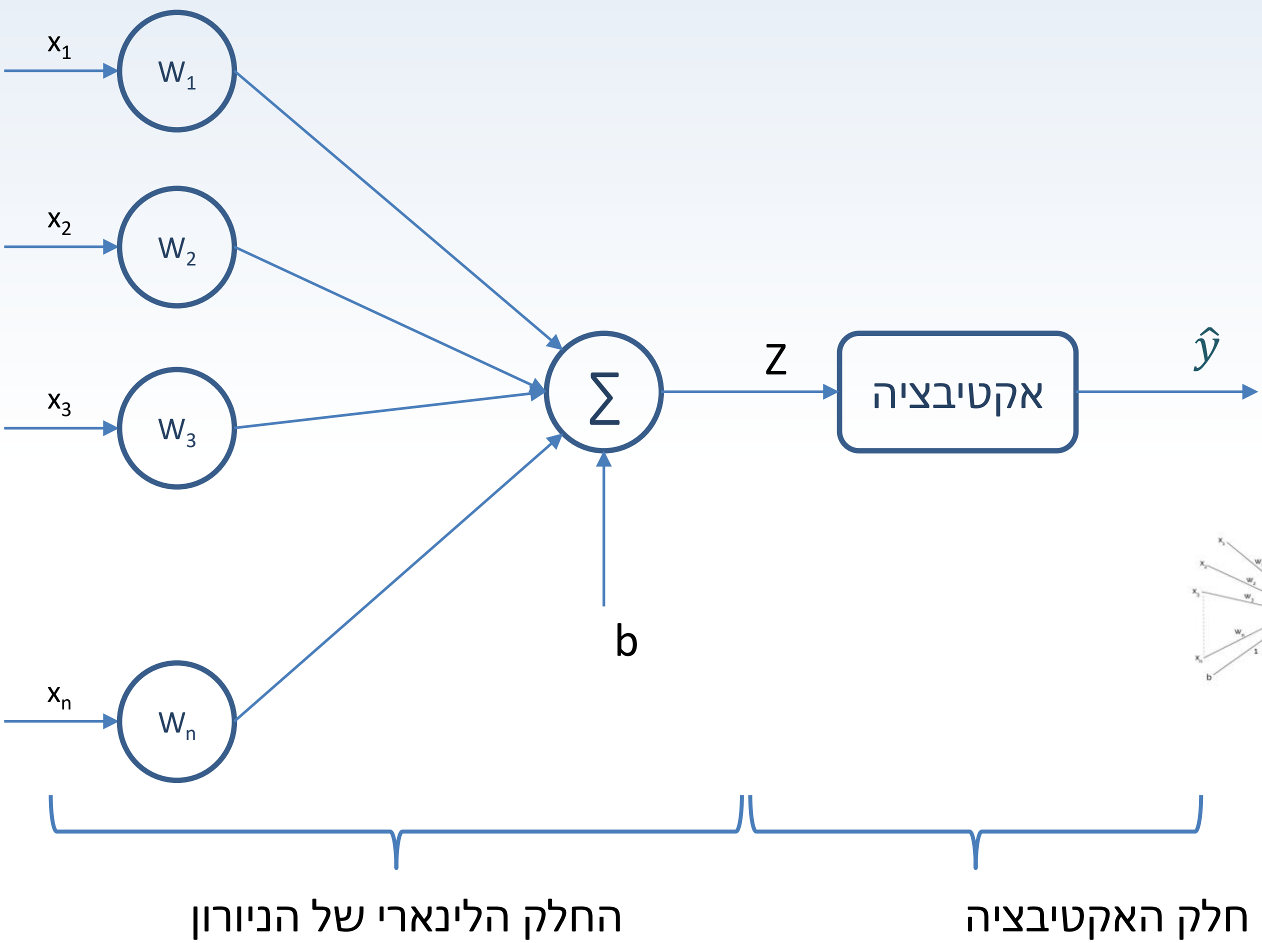


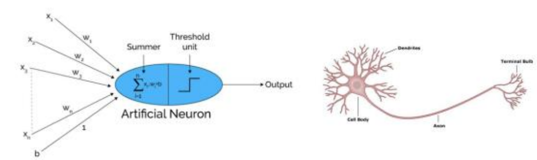
פונקציה לוגיסטית - משמעות



אם z = 0
ערך הפונקציה שווה ל-0.5.
אם z חיובי מאד – ערך הפונקציה שואף ל-1.
אם z שלילי מאד – ערך הפונקציה שואף ל-0.
בין לבין, ערכי הפונקציה רציפים
קיבלנו את הפונקציה הלוגיסטית
פונקציה רציפה בין 0 ל-1,
הניתנת לגזירה

ארכיטקטורה של נוירון בודד

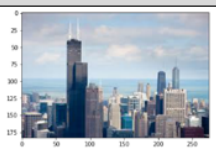
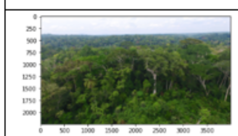
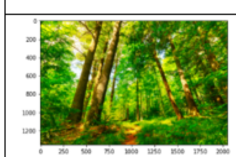




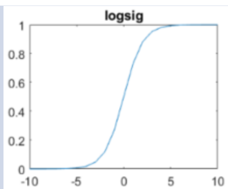
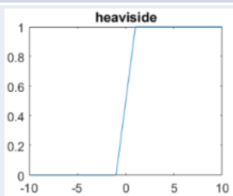
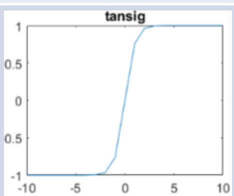
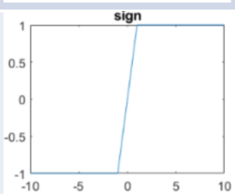
תרגיל 1

סיווג על ידי רגרסיה לוגיסטית

- נתונות תמונות בעלות שני מאפיינים:
  - x1 ממוצע האדום בתמונה
  - x2 ממוצע הכחול בתמונה
- נתון מסווג מסוג רגרסיה לוגיסטית בעל הפרמטרים הבאים:
  - w1=0.06
  - w2=0.04
  - b=-10

Image	X_1 (red)	X_2 (blue)	$z = w_1 \cdot x_1 + w_2 \cdot x_2 + b$	$\hat{y} = \sigma(z)$	סוג יער=0 עיר=1
	153	173			
	93	83			
	81	124			
	90	18			

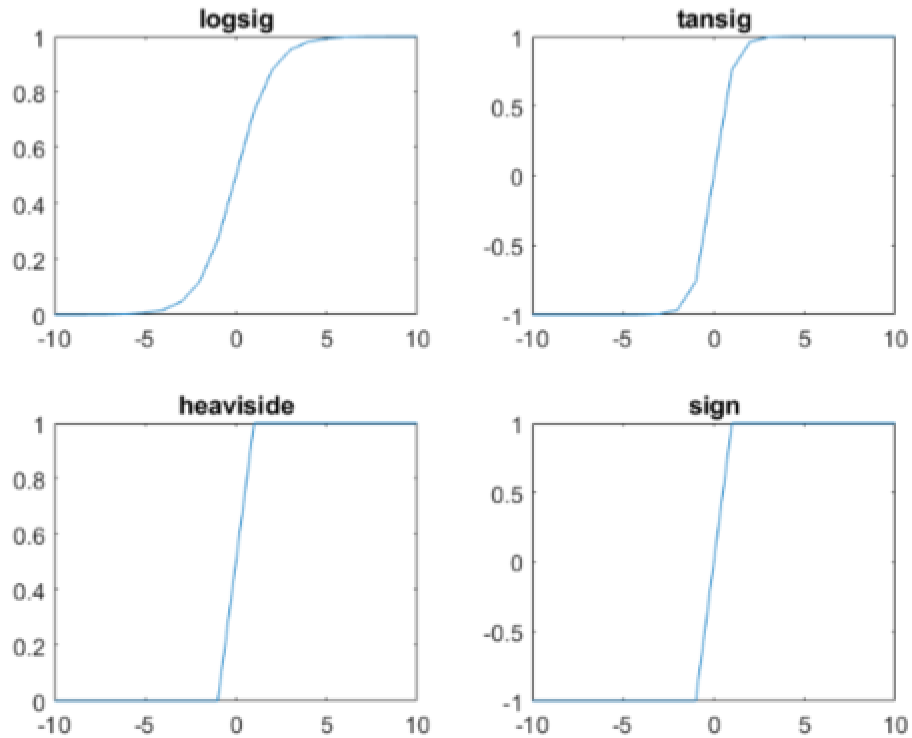
פונקציות טווחים ב PYTHON

שם	טווח	רציפה	ויזואליציה
לוגיסטית	0 עד 1	כן	
מדרגה	0 עד 1	לא	
טנגנס היפרבולי	-1 עד 1	כן	
סימן	-1 עד 1	לא	

תרגיל 2
פונקציות טווחים ב PYTHON

(מחברת Colab):

הציגו על המסך 4 גרפים שמייצגים את פונקציות הטווחים
צרו וקטור ערכים לציר ה X מ -10 עד 10 במרווחים של 1



סיווג לשתי מחלקות

משמעות הפלט של הנוירון בפונקציה לוגיסטית

הערכים בפלט: כל ערך בין 0 ל 1

נניח שיש לנו בעיית סיווג:

נתונים של קוטר האבטיח. וכתשובה האם הפרי בשל או לא

- ✓ אם הפלט 1 – הפרי בשל, אם הפלט 0 – הפרי לא בשל
- ✓ המודל סיווג 0.83 – מה המשמעות?

0.83 אומר שיש סיכוי של 83% שהפרי בשל, ושל 17% שהפרי אינו בשל.

- ✓ המודל סיווג 0.83 – האם הפרי בשל?

מעל 0.5 – נקבע שהפרי בשל, מתחת ל 0.5 נקבע שהפרי לא בשל.

תרגול

אימנת מודל סיווג שמנבא האם תלמיד יסיים תואר ראשון על סמך נתוני הבחינה הפסיכומטרית

הפלט עבור דוגמא חדשה חזה h=0.7

אילו מההגדים נכונים?

A ❑ 70% שהתלמיד יסיים תואר I

B ❑ 30% שהתלמיד יסיים תואר I

C ❑ 100% שהתלמיד יסיים תואר I

D ❑ 30% שהתלמיד לא יסיים תואר ראשון

תרגול - פתרון

אימנת מודל סיווג שמנבא האם תלמיד יסיים תואר ראשון על סמך נתוני הבחינה הפסיכומטרית

הפלט עבור דוגמא חדשה חזה h=0.7

אילו מההגדים נכונים?

A ❑ 70% שהתלמיד יסיים תואר I נכון

B ❑ 30% שהתלמיד יסיים תואר I

C ❑ 100% שהתלמיד יסיים תואר I

D ❑ 30% שהתלמיד לא יסיים תואר ראשון נכון

משמעות הפלט של הנוירון בפונקציה לוגיסטית

ערך הפונקציה הלוגיסטית שווה ל-50% כאשר הקלט המשוקלל שווה ל-0.

לכן:

אם **z** גדול שווה מ 0 – נסווג למחלקה אחת (y=1)

אם **z** קטן מ 0 – נסווג למחלקה אחרת (y=0)

קו ההחלטה (decision boundary-ה) הוא Z = $\sum_{i=1}^{n} \mathrm{w(i)x(i)} + b = 0$

אם $Z \geq 0$, y = 1

אם $Z < 0$, y = 0

$w = [-1 \quad 0\ ]$, b = 5

$y = 1\ if\ (-1)x_1 + 0x_2 + 5 \geq 0$

$5 - x_1 \geq 0$

-$x_1 \geq -5$

$x_1 \leq 5$

קיבלנו קו הפרדה ליניארי.

אימנת מודל סיווג שמנבא האם תלמיד יסיים תואר ראשון על סמך נתוני הבחינה הפסיכומטרית x1 ונתוני ציוני הבגרות x2:

$$\hat{y} = \frac{1}{1+e^{-z}} \quad , z = w1x1 + w2x2 + b$$

המשקולות של המודל שלך הם:

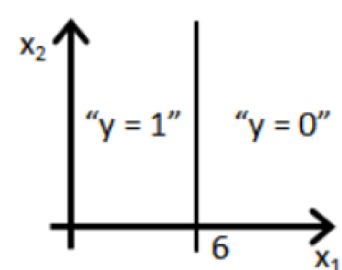
0 = w1

-1 = w2

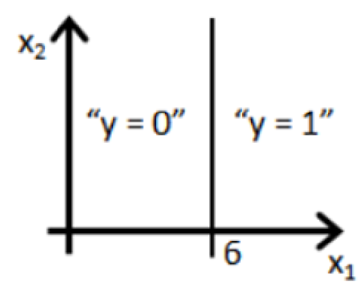
6 = b

איזה גרף מציין נכון את קו ההפרדה של הנתונים:

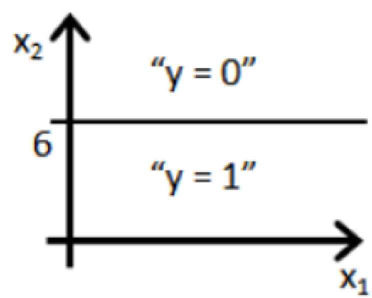
A



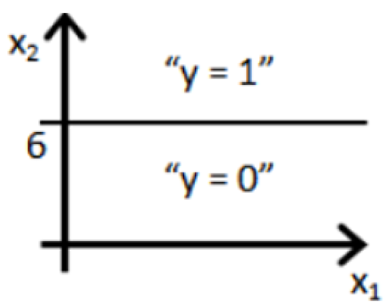
B



C



D



תרגול - פתרון

אימנת מודל סיווג שמנבא האם תלמיד יסיים תואר ראשון על סמך נתוני הבחינה הפסיכומטרית x1 ונתוני ציוני הבגרות x2:

$$\hat{y} = \frac{1}{1+e^{-z}} \quad , z = w1x1 + w2x2 + b$$

המשקולות של המודל שלך הם:

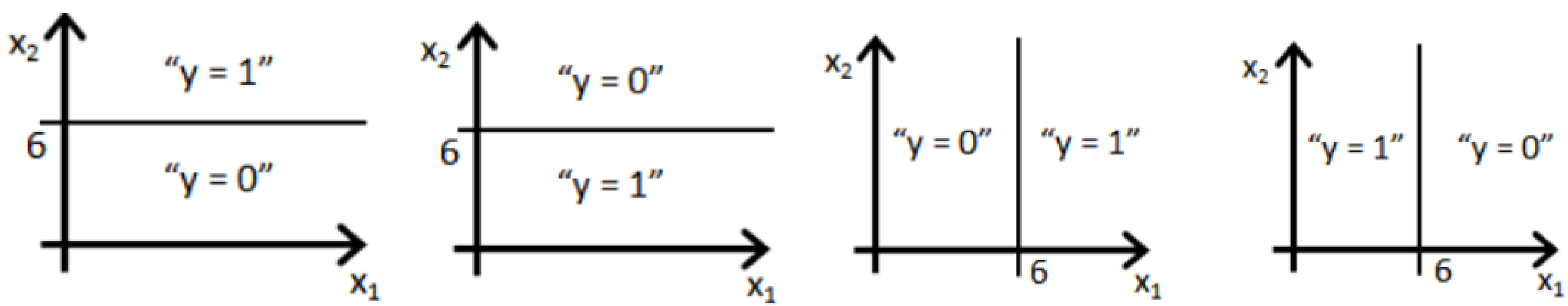
0 = w1

-1 = w2

6 = b

איזה גרף מציין נכון את קו ההפרדה של הנתונים:

D C B A

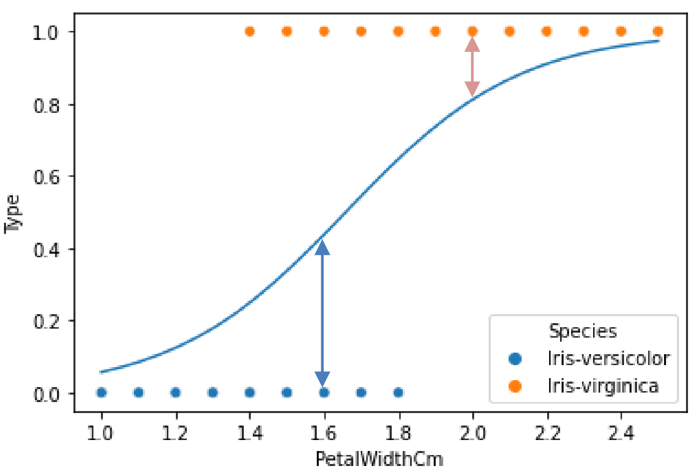


Cost function עבור רגרסיה לוגיסטית

פונקציית המחיר - BINARY CROSS ENTROPY

$$l^{(i)} = -y^{(i)} \log \hat{y}^{(i)} - (1 - y^{(i)}) \log(1 - \hat{y}^{(i)})$$

$$J = \frac{1}{m} \sum_{i=1}^{m} l^{(i)}$$



$\hat{y}$(i) – הערך החזוי בכל דגימה

$$\hat{y}^{(i)} = \left(\frac{1}{1+e^{-z(i)}}\right)$$

y(i) – הערך האמיתי בכל דגימה

m – מספר הדגימות

i – אינדקס על כל הדגימות

תרגיל 3

חישוב פונקציית מחיר

נתונות הדוגמאות הבאות. חשבו את פונקציית המחיר על כל אחת מהדוגמאות ואת המחיר הכולל J

i	$y^{(i)}$	$\hat{y}^{(i)}$	$-y^{(i)}\log\hat{y}^{(i)}-(1-y^{(i)})\log(1-\hat{y}^{(i)})$
1	0	0.1	
2	0	0.5	
3	0	0.6	
4	1	0.9	
5	1	0.5	
6	1	0.4	

$$J=\frac{1}{6}\sum_{i=1}^{6} l^{(i)}=?$$

המטרה – למזער את הטעות עבור ערכי המשקולות השונים

$$\hat{y} = \frac{1}{1+e^{-z}} \quad \begin{cases} if\,\hat{y} \geq 0.5, y\_predicted = 1 \\ if\,\hat{y} < 0.5, y\_predicted = 0 \end{cases}$$

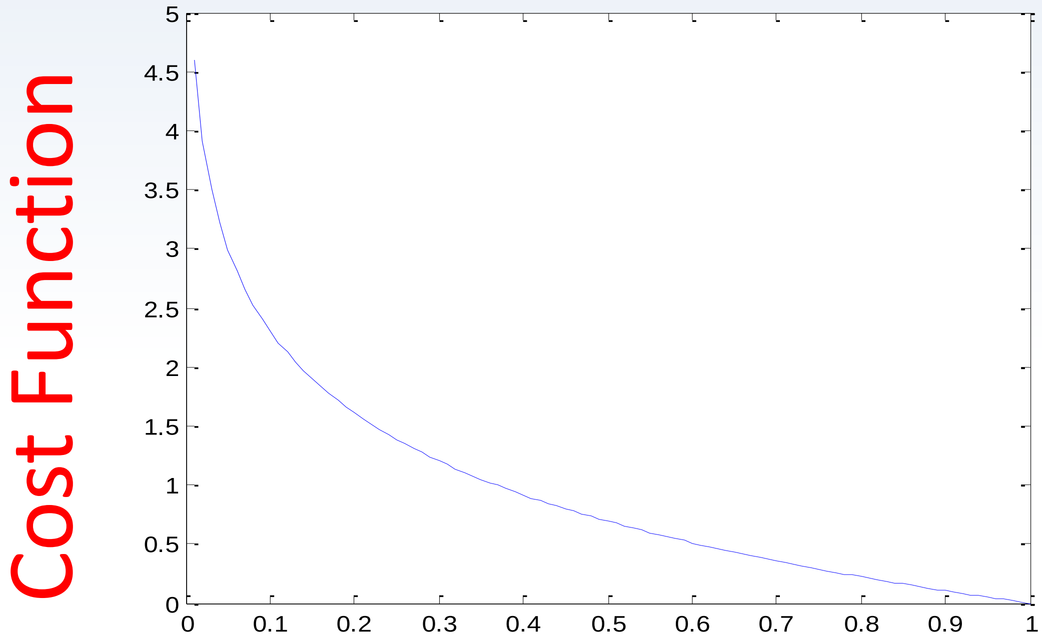
Cost Function

$$J = \begin{cases} -\log(\hat{y}) \quad if\ y = 1 \\ -\log(1-\hat{y}) \quad if\ y = 0 \end{cases}$$

y הנתונים האמיתיים

פונקציית העלות עבור Y=1

$$-\log(\hat{y})$$



$$0 < \hat{y} < 1$$

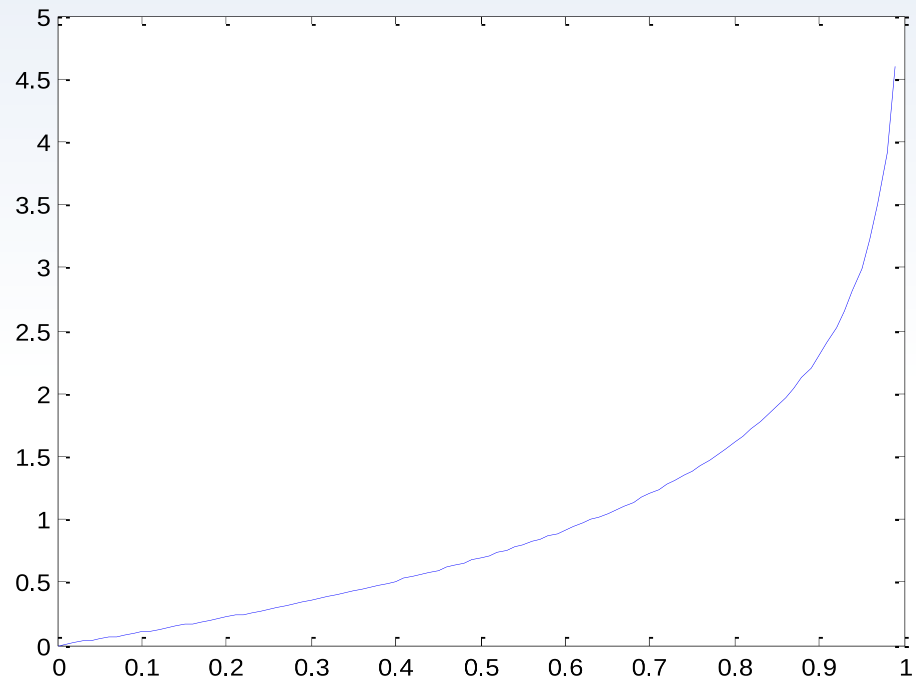
כאשר y=1 וגם הפלט הוא 1 – ה-cost הוא 0.

כאשר y=1 אבל הפלט הוא 0 – ה-cost שואף לאינסוף

פונקציית העלות עבור Y=0

$$-\log(1-h(w,x))$$

Cost Function



$$0<h(w,x)<1$$

כאשר y=0 וגם הפלט הוא 0 – ה-cost הוא 0.

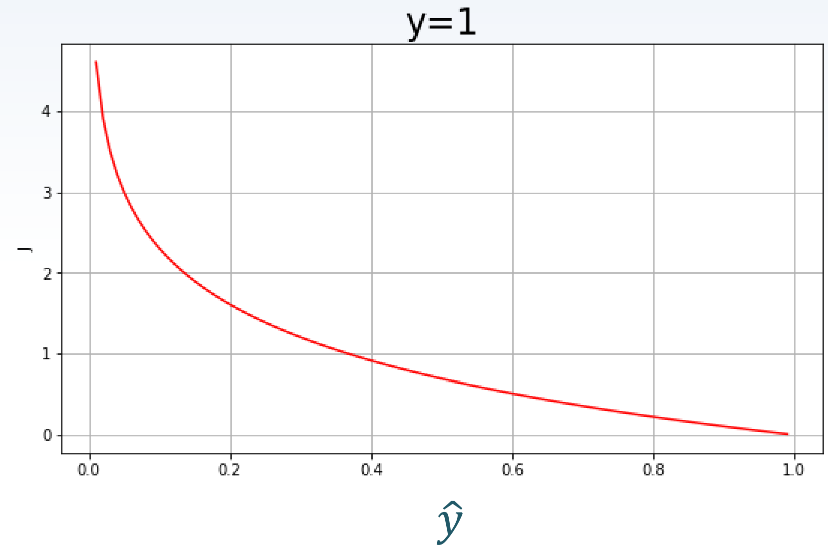
כאשר y=0 אבל הפלט הוא 1 – ה-cost שואף לאינסוף

תרגיל 4
ויזואליזציה של פונקציית המחיר

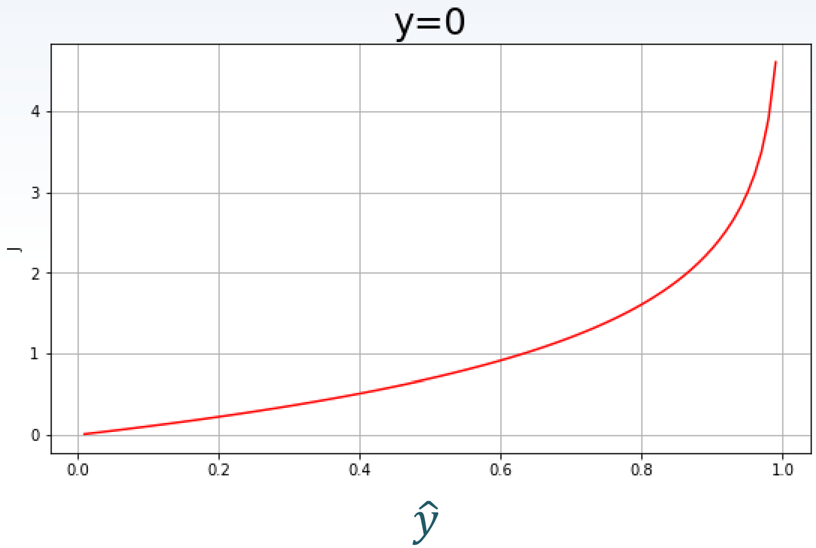
- שרטטו את פונקציית המחיר עבור דוגמאות מסוג "0" שהחיזוי עבורן בתחום 0.0-0.99
- שרטטו את פונקציית המחיר עבור דוגמאות מסוג "1" שהחיזוי עבורן בתחום 0.01-1.0

ויזואליזציה של פונקציית המחיר

$$J = -\log(\hat{y})$$



$$J = -\log\left(1 - \hat{y}^{(i)}\right)$$



אלגוריתם GRADIENT DESCENT

- הגדרנו את פונקציית המחיר

$$J(w,b)$$

- נאתחל:

$$w \leftarrow 0, b \leftarrow 0$$

- נתקדם לכיוון המינימום:

$$w \leftarrow w - \alpha \frac{\partial J(w,b)}{\partial w}$$

$$b \leftarrow b - \alpha \frac{\partial J(w,b)}{\partial b}$$

אלגוריתם GRADIENT DESCENT לאימון רגרסיה לוגיסטית

- הגדרנו את פונקציית המחיר

$$J(w, b)$$

- נאתחל:

$$w \leftarrow 0, b \leftarrow 0$$

- נתקדם לכיוון המינימום:

$$w \leftarrow w - \alpha \frac{1}{m} \sum_{i=1}^{m} \overbrace{\left(\hat{y}^{(i)} - y^{(i)}\right)}^{\text{השגיאה בין הערך החזוי לאמיתי}} x^{(i)}$$

$$b \leftarrow b - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)}\right)$$

שאלות למחשבה:

- מהם ההבדלים בין נוסחאות האימון של רגרסיה לוגיסטית לרגרסיה ליניארית?
- מהם ההבדלים בין נוסחאות האימון של רגרסיה לוגיסטית לפרספטרון?

gradient descent עבור רגרסיה לוגיסטית

האם GRADIENT DESCENT עבור רגרסיה ליניארית ועבור רגרסיה לוגיסטית זהה?

$$w(j) = w(j) - (\alpha/m)\sum_{i=1}^{m}\left(\hat{y}^{(i)} - y^{(i)}\right)x(j)(i)$$

ברגרסיה ליניארית:

$$\hat{y}^{(i)} = \sum_{j=1}^{n} w(j)x(j) + b$$

ברגרסיה לוגיסטית:

$$\hat{y}^{(i)} = \frac{1}{1+e^{-\sum_{j=1}^{n} w(j)x(j)+b}}$$

הערכת ביצועים עבור רגרסיה לוגיסטית

חישוב אחוז הדגימות שהמודל צדק בהם

y - וקטור הערכים האמיתיים, ערכים בדידים 0 או 1

0	0	0	0	1	1	1	1	1	1

$\hat{y}$ - וקטור הערכים החזויים, ערכים בין 0 ל 1

0.2	0.1	0.3	0.6	0.8	0.6	0.4	0.9	0.8	0.7

if $\hat{y}$ >=0.5, p=1
if $\hat{y}$ <0.5, p =0

p - וקטור הערכים החזויים, ערכים בדידים 0 או 1

0	0	0	1	1	1	0	1	1	1

```
Accuracy = all correct samples / all samples = 8/10, 80%
```

רגרסיה לוגיסטית ב SKLEARN

- # ייבוא ספריה
- `from sklearn.linear_model import LogisticRegression`
- # בחירת מאפיינים
- `X=df[[…]].to_numpy()`
- # בחירת תנאי לוגי
- `y=1*(df[…]==…).to_numpy()`
- # הגדרת מסווג
- `clf = LogisticRegression()`
- # אימון
- `clf.fit(X, y)`