# תרגיל 1 - יותם גרדוש - 208541334

## שאלות הקדמה

נתון הפרספטרון הבינארי $y = H(\bar{w} \cdot \bar{x})$, שמקבל דוגמאות דו־מימדיות $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

1. אם ידוע כי $w_2 = 0$ וכי הדוגמה $\bar{x} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$ מסווגת ב־ 0, מה ניתן לומר על הסיווג של הדוגמה $\bar{x} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$?

(א) תסווג כ־ 1

(ב) תסווג כ־ 0

(ג) לא ניתן לדעת

מנתון: $w_2 = 0 \Rightarrow w_1x_1 + 0x_2 = 0 \Leftrightarrow w_1x_1 = 0$

$w_1 < 0 \Leftarrow w_1 \cdot 7 < 0 \Leftrightarrow H\left(\bar{w}^T \begin{bmatrix} 7 \\ 2 \end{bmatrix}\right) = 0$

מהמהלך האלגברי הנ"ל ניתן ללמוד ש-$w_1<0$ לכן עבור $X = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$ ; $\bar{w} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ נקבל כי $\bar{w}\cdot\bar{x} < 0$ לכן כנ' הפרס"ט תסווג את $\bar{x}$ ל-0.

2. עבור $\bar{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, מה מהבאים **יבטיח** שדוגמה תסווג ב־ 1? (תיתכן יותר מתשובה אחת נכונה)

(א) $x_1 = x_2$

(ב) $x_1 > x_2$

(ג) $x_1 > -x_2$

(ד) $x_2 < -x_1$

(ה) $x_2 > -x_1$

(ו) $-x_1 > -x_2$

$H(\bar{w}^T\bar{x}) = 1 \Leftrightarrow \bar{w}^T\bar{x} > 0 \Leftrightarrow w_1x_1 + w_2x_2 > 0 \Leftrightarrow x_1 + x_2 > 0 \Leftrightarrow x_1 > -x_2 \lor x_2 > -x_1$

ניתן לראות שזהו קו גבולי בין 2 הפתרונות המסומנים לבין הסיווג הרצוי

3. נתון פרספטרון בינארי עם וקטור משקולות $\bar{w}$.
נתונות שתי דוגמאות $\bar{x}^1$ ו־$\bar{x}^2$, עבורן מתקיים:

$$\bar{w}^T\bar{x}^1 \neq 0 \ , \ \bar{w}^T\bar{x}^2 \neq 0$$

מה מהבאים **יבטיח** ששתי דוגמאות **לא** יסווגו באותו הסיווג?

(א) הדוגמאות מקיימות $\bar{x}^1 = \alpha\bar{w}$ עבור $\alpha > 0$ ו־ $\bar{x}^2 = \beta\bar{w}$ עבור $\beta > 0$

(ב) הדוגמאות מקיימות $\alpha\left(\bar{x}^1 + \bar{x}^2\right) = \bar{w}$ עבור $\alpha$ סקלר כלשהו

(ג) הדוגמאות מקיימות $\bar{w}^T\bar{x}^1 = \alpha\bar{w}^T\bar{x}^2$ עבור $\alpha < 0$

(ד) כל התשובות לא מבטיחות ששתי הדוגמאות לא יסווגו באותו הסיווג

כך $\bar{w}^T\bar{x}^1 = \alpha\bar{w}^T\bar{x}^2$ עבור $\alpha < 0$ כלשהו מבטיח סיווגים הפוכים. זאת מכיוון ששיוויון בין המשוואות מבטיח גם שיוויון בסימן. מכיוון ש-$\alpha<0$ נוכל להסיק כי הסימן של $\bar{w}^T\bar{x}^1$ ושל $\bar{w}^T\bar{x}^2$ הפוכים מה שמבטיח לנו סיווגים הפוכים לפי הגדרת פונ' המדרגה H.

# שאלה 1

עבור הפרספטרון הבינארי: $y = H(\bar{w}\cdot\bar{x})$, נתונות הדוגמאות הבאות וסיווגן הרצוי:

- $\mathbf{x}^1 = (2,2)$, $y^1 = 1$
- $\mathbf{x}^2 = (1,3)$, $y^2 = 1$
- $\mathbf{x}^3 = (-1,0)$, $y^3 = 0$
- $\mathbf{x}^4 = (-1,2)$, $y^4 = 0$

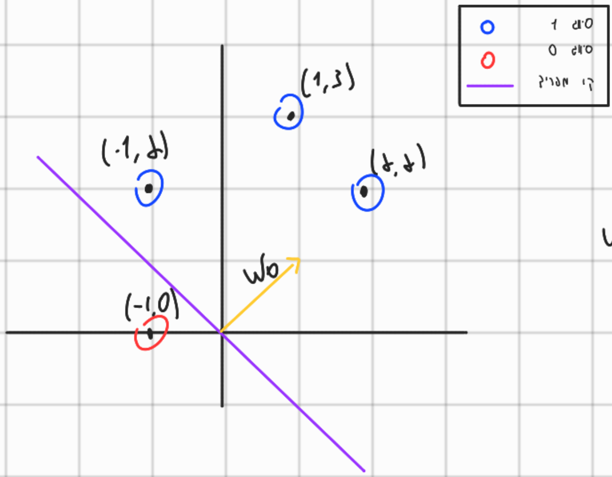
## סעיף א'

התחילו מוקטור משקולות $\mathbf{w} = [1,1]$. ציירו באופן סכמטי את הנקודות, את וקטור המשקולות ואת הישר המפריד.

הציגו לפרספטרון את הדוגמאות הנתונות לפי הסדר, ועדכנו את המשקולות בהתאם לכלל הלמידה של הפרספטרון עד להתכנסות (כלומר, עד שהפרספטרון מסווג נכונה את כל ארבע הדוגמאות).

ציירו באופן סכמטי את הנקודות, את וקטור המשקולות ואת הישר המפריד לאחר ההתכנסות.

ציור סכמטי של הנקודות.




חישוב נוסחת הישר המפריד.

$$w_1x_1 + w_2x_2 = 0 \Rightarrow w_2x_2 = -w_1x_1 \Rightarrow x_2 = -x_1\frac{w_1}{w_2}$$

איטרציה ראשונה -

$x_1$, $x_2$ ו-$x_3$ מסווגים נכון. $x_4$ לא לכן נעדכן $w$:

$$w_1 = w_0 + (2\cdot 0-1)\cdot x_4 = \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}2\\-1\end{pmatrix}$$

איטרציה שניה -

$$x^1:\ H(w_1x^1) = H(4-2) = 1 = y^1,$$

$$x^2:\ H(w_1x^2) = H(2-3) = 0 \neq y^2 \xRightarrow{\text{עדכון } w} w_2 = w_1 + (2\cdot 1-1)\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}2\\-1\end{pmatrix} + \begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix}$$

$$x^3:\ H(w_2x^3) = H(-3) = 0 = y^3$$

$$x^4:\ H(w_2x^4) = H(-3+4) = 1 \neq y^4 \xRightarrow{\text{עדכון } w} w_3 = w_2 + (2\cdot 0-1)\begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix} - \begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}4\\0\end{pmatrix}$$

## אינרציה שלישית

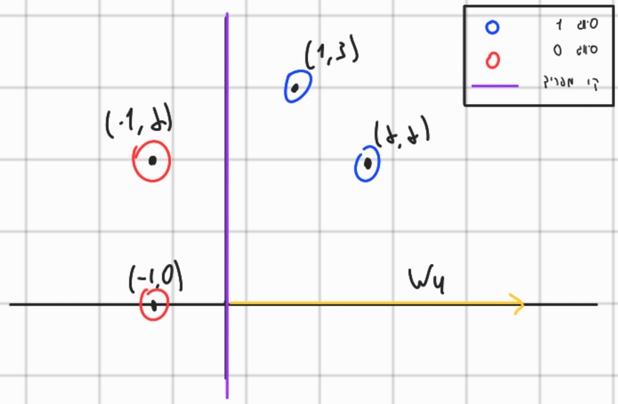
$x^1: H(w_3x^1) = H(2) = 1 = y^1$

$x^2: H(w_3x^2) = H(4) = 1 = y^2$

$x^3: H(w_3x^3) = H(-1) = 0 = y^3$

$x^4: H(w_3x^4) = H(-1) = 0 = y^4$

שלב לאחר התכנסות =>




**סעיף ב'**

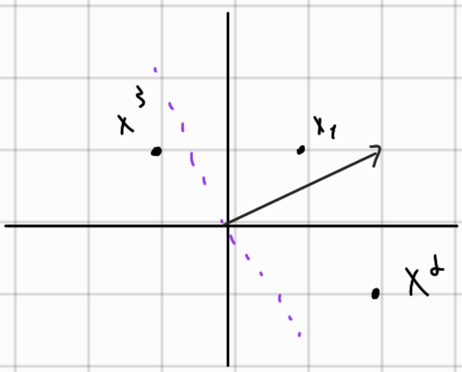
נתון פרספטרון בינארי עם סף $y = \mathrm{H}(w_1x_1 + w_2x_2 + T)$.

מצאו ערכי $\bar{w}$ ו- $T$ עבור מתקיימים הכללים הבאים. הסבירו את תשובתכם.

1. $y = 1$ אם ורק אם $2x_1 + x_2 > 0$
2. $y = 1$ אם ורק אם $x_1 < 3x_2 + 4$

1. $y = 1 \iff 2x_1 + x_2 > 0$

סתם למען אינטואיציה $x^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $x^2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ $x_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$




אבחר $b = 0$ $w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

כיוון ראשון -

$2x_1 + x_2 > 0 \Rightarrow y = 1$:

$$H(w^Tx+b) = H(w_1x_1 + w_2x_2 + b) \overset{\text{הצבת } w,b}{=} H(2x_1 + x_2 + 0) \overset{\text{מהנחה } 2x_1+x_2>0}{=} 1$$

כיוון שני -

$y = 1 \Rightarrow 2x_1 + x_2 > 0$:

$$y = 1 \Rightarrow H(w^Tx+b) = 1 \overset{\text{מהגדרת } H}{\Rightarrow} w_1x_1 + w_2x_2 + b > 0 \overset{\text{הצבת } w,b}{=} 2x_1 + x_2$$

לפי ההגדרה $w = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ ! $b = 0$ קיבלנו בחירה זו כיוונית נכונה ∎

ג. $y=1 \iff x_1<3x_2+4$

קודם ראשון ארצה לסדר את הביטוי מצד ימין של הגדרה: $-x_1+3x_2+4>0 \iff x_1<3x_2+4$

אבחר $w=\begin{pmatrix}-1\\3\end{pmatrix}$ $b=4$:

$$y=1 \iff H(w^Tx+b)=1 \overset{\text{הגדרת } H}{\iff} w^Tx+b>0 \iff w_1x_1+w_2x_2+b>0 \overset{\text{הגדרת } w,b}{\iff} -x_1+3x_2+4>0$$

קיבלנו שרשרת של גרירות דו כיווניות רק שעבור $w=\begin{pmatrix}-1\\3\end{pmatrix}$, $b=4$ מתקיים

$$y=1 \iff -x_1+3x_2+4>0 \iff 3x_2+4>x_1 \quad \blacksquare$$

**סעיף ג'**

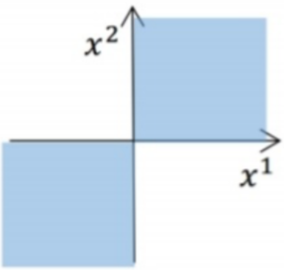
ראינו בכיתה כי פרספטרון בינארי לא מסוגל לבצע הפרדות שאינן לינאריות. בשביל הפרדות כאלו, נוכל להשתמש ברשת של כמה פרספטרונים בינאריים.

בנו רשת של פרספטרונים שתממש את הפונקציה **NXOR**:

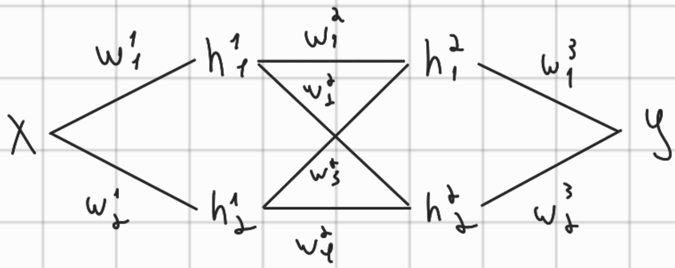
$y=1$ אם ורק אם ($x_1>0$ וגם $x_2>0$) <u>או</u> ($x_1<0$ וגם $x_2<0$)

ראו בציור המצורף מטה.

השתדלו לבנות את הרשת היעילה ביותר שניתן לבנות, כלומר, עם כמה שפחות תאים/נוירונים (רמז: בשכבה הראשונה לא אמורים להיות יותר משלושה).




נגדיר את הרשת הבאה




נניח בנוסף כי: $H(0)=0$

כך ש:

$$x=\begin{pmatrix}x_1\\x_2\end{pmatrix},\quad w_1=\begin{pmatrix}1\\0\end{pmatrix},\quad w_2=\begin{pmatrix}0\\-1\end{pmatrix} \Rightarrow h_1^1=H(x_1),\quad h_2^1=H(-x_2),$$

$$w_1^2=w_4^2=1,\quad w_2^2=w_3^2=-2 \Rightarrow h_1^2=H(h_1^1-2h_2^1),\quad h_2^2=H(-2h_1^1+h_2^1),$$

$$w_1^3=1,\quad w_2^3=1 \Rightarrow y=H(h_1^2+h_2^2)$$

הסבר מילולי -

מטרת הרשת היא לזהות שבשני הקורדינאטות של x הם שווי סימן.

כאשר הם שווי סימן הביטויים $h_1^1$ ו $h_2^1$ יהיו אחד עם ערך 0 והשני 1 (תלוי בסימן)

לכן אחד מבין $h_1^2$ או $h_2^2$ בוודאות יהיה 1. מה שמבטיח $y=1$ כנדרש במקרה זה.

כאשר הם בסימנים הפוכים: אם שלילי $x_1=$ ו חיובי $x_2=$ נקבל $h_1^1=h_2^1=0 \Rightarrow h_1^2=h_2^2=0 \Rightarrow y=0$

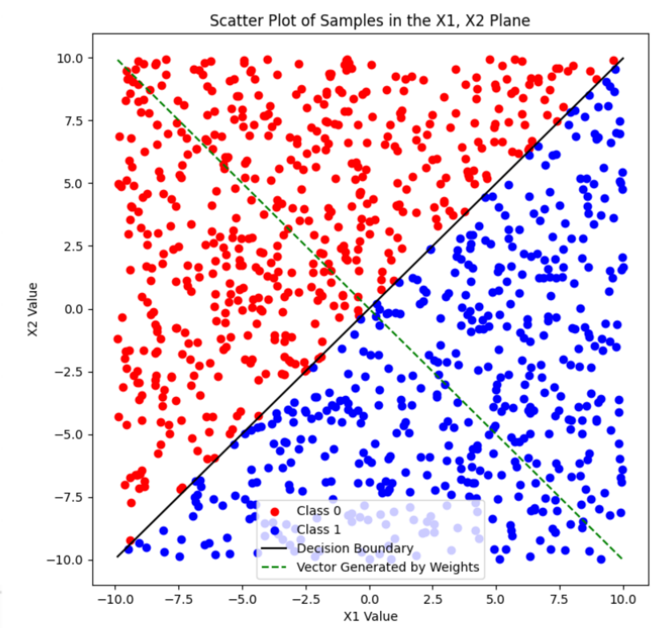
אם חיובי $x_1=$ ו שלילי $x_2=$ הבחירה של $w_2^2=w_3^2=-2$ מבטיחה

לכן כי בחישוב $h_1^2, h_2^2$ נקבל: $h_1^2=h_2^2=0 \Rightarrow y=0$ כנדרש

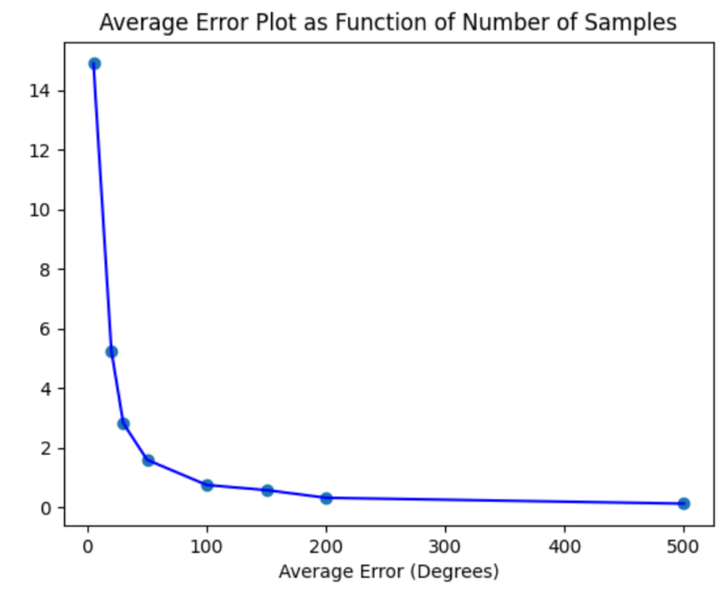
לכן קיבלנו רשת שמממשת את NXOR ☒

# שאלה 2

.3




.4




4 - הסבר אינטואיטיבי:

$w^*$ אופטימלי הוא הוקטור $w^*=\binom{1}{-1}$ כאשר מכיוון שלפי ההגדרה בסעיף (ג) כי $y=1 \iff x_1>x_2$
נקבל כי $y=H(\bar{w}^T\bar{x})=H(x_1-x_2)$. לכן $[y=1 \iff x_1-x_2>0 = x_1>x_2]$ וגם $[y=0 \iff x_1-x_2<0 = x_1<x_2]$
אם כן נקבל כי כל הנקודות יסווגו נכון עבור $w^*$ הנ"ל

$w^*=\binom{1}{-1}$ הוא לא הוקטור האופטימלי היחיד למשל הקבוצה האין סופית $S=\{\binom{a}{-a} \mid a\in R\}$
היא קבוצה של וקטורים אופטימליים.

הסבר גרף -
הגרף מראה את ההפרש בזווית מוחלט בממוצע בין $\bar{w}$ משקולות האלגוריתם לבין $w^*$ אופטימלי.
מעצם הגדרת $w^*$ הוא מסווג נכון כל דוגמא במרחב הנתון לכן ככל שנקבל יותר דוגמאות,
משקולות אלגוריתם הפרספטרון שואפות למשקולות האופטימליות. זה מתבטא בגרף ע"י מגמת הירידה
בערך השגיאה, באופן תואם אם מספר הדוגמאות P שואף ל-$\infty$ אז ערך השגיאה שואף ל-0,
וזה שקול לכך שמשקלות המודל הן אופטימליים.

# חישוביות וקוגניציה - תרגיל 1 ־ פרסנפטרון בינארי

להגשה עד: 18/01/2024 בשעה 21:00

שימו לב: בתחילת התרגיל מופיעות כמה שאלות הקדמה אמריקאיות. יש להגיש את התשובות אליהן, עם משפט נימוק קצר לכל שאלה. לאחר מכן, שאלה 1 היא שאלה אנליטית ושאלה 2 היא שאלת תכנות.
לשימושכם, אלגוריתם הלמידה של הפרספטרון מופיע בסוף התרגיל.

## שאלות הקדמה

נתון הפרספטרון הבינארי $y = H(\bar{w} \cdot \bar{x})$, שמקבל דוגמאות דו־מימדיות $\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

1. אם ידוע כי $w_2 = 0$ וכי הדוגמה $\bar{x} = \begin{bmatrix} 7 \\ 2 \end{bmatrix}$ מסווגת ב־ 0, מה ניתן לומר על הסיווג של הדוגמה $\bar{x} = \begin{bmatrix} 5 \\ -10 \end{bmatrix}$?
   (א) תסווג כ־ 1
   (ב) תסווג כ־ 0
   (ג) לא ניתן לדעת

2. עבור $\bar{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, מה מהבאים **יבטיח** שדוגמה תסווג ב־ 1? (תיתכן יותר מתשובה אחת נכונה)
   (א) $x_1 = x_2$
   (ב) $x_1 > x_2$
   (ג) $x_1 > -x_2$
   (ד) $x_2 < -x_1$
   (ה) $x_2 > -x_1$
   (ו) $-x_1 > -x_2$

3. נתון פרספטרון בינארי עם וקטור משקולות $\bar{w}$.
   נתונות שתי דוגמאות $\bar{x}^1$ ו־$\bar{x}^2$, עבורן מתקיים:

$$\bar{w}^T\bar{x}^1 \neq 0 \ , \ \ \bar{w}^T\bar{x}^2 \neq 0$$

   מה מהבאים **יבטיח** ששתי דוגמאות **לא** יסווגו באותו הסיווג?
   (א) הדוגמאות מקיימות $\bar{x}^1 = \alpha\bar{w}$ עבור $\alpha > 0$ ו־ $\bar{x}^2 = \beta\bar{w}$ עבור $\beta > 0$
   (ב) הדוגמאות מקיימות $\alpha\left(\bar{x}^1 + \bar{x}^2\right) = \bar{w}$ עבור $\alpha$ סקלר כלשהו
   (ג) הדוגמאות מקיימות $\bar{w}^T\bar{x}^1 = \alpha\bar{w}^T\bar{x}^2$ עבור $\alpha < 0$
   (ד) כל התשובות לא מבטיחות ששתי הדוגמאות לא יסווגו באותו הסיווג

## שאלה 1

עבור הפרספטרון הבינארי: $y = H(\bar{w} \cdot \bar{x})$, נתונות הדוגמאות הבאות וסיווגן הרצוי:

- $\mathbf{x}^1 = (2,2)$, $y^1 = 1$
- $\mathbf{x}^2 = (1,3)$, $y^2 = 1$
- $\mathbf{x}^3 = (-1,0)$, $y^3 = 0$
- $\mathbf{x}^4 = (-1,2)$, $y^4 = 0$

### סעיף א'

התחילו מוקטור משקולות $\mathbf{w} = [1,1]$. ציירו באופן סכמטי את הנקודות, את וקטור המשקולות ואת הישר המפריד.

הציגו לפרספטרון את הדוגמאות הנתונות לפי הסדר, ועדכנו את המשקולות בהתאם לכלל הלמידה של הפרספטרון עד להתכנסות (כלומר, עד שהפרספטרון מסווג נכונה את כל ארבע הדוגמאות).

ציירו באופן סכמטי את הנקודות, את וקטור המשקולות ואת הישר המפריד לאחר ההתכנסות.

### סעיף ב'

נתון פרספטרון בינארי עם סף $y = \mathrm{H}\left(w_1 x_1 + w_2 x_2 + T\right)$.

מצאו ערכי $\bar{w}$ ו- $T$ עבורם מתקיימים הכללים הבאים. הסבירו את תשובתכם.

1. $y = 1$ אם ורק אם $2x_1 + x_2 > 0$
2. $y = 1$ אם ורק אם $x_1 < 3x_2 + 4$

### סעיף ג'

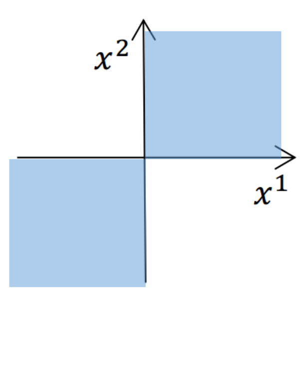
ראינו בכיתה כי פרספטרון בינארי לא מסוגל לבצע הפרדות שאינן לינאריות. בשביל הפרדות כאלו, נוכל להשתמש ברשת של כמה פרספטרונים בינאריים.

בנו רשת של פרספטרונים שתממש את הפונקציה **NXOR**:

$y = 1$ אם ורק אם ($x_1 > 0$ וגם $x_2 > 0$) <u>או</u> ($x_1 < 0$ וגם $x_2 < 0$)

ראו בציור המצורף מטה.

השתדלו לבנות את הרשת היעילה ביותר שניתן לבנות, כלומר, עם כמה שפחות תאים/נוירונים (רמז: בשכבה הראשונה לא אמורים להיות יותר משלושה).



## שאלה 2

1. כתבו פונקציה שמקבלת כקלט מטריצה בגודל $N \times P$, בה יש $P$ דוגמאות ממרחב $N$־מימדי. הפונקציה מקבלת גם וקטור באורך $P$, בו כל התיוגים הרצויים לדוגמאות.
בעזרת אלגוריתם הלמידה של הפרספטרון, הפונקציה צריכה למצוא וקטור משקולות $\bar{w}$, שבעזרתו הפרספטרון מסווג נכונה את כל הדוגמאות.

2. הגדירו $N = 2$, והגרילו $P = 1000$ נקודות דו־מימדיות, כאשר כל קואורדינטה $(x_1, x_2)$ מתפלגת בהתפלגות אחידה (רציפה) בין $-10$ ל־ $10$. שתי הקואורדינטות מוגרלות באופן בלתי תלוי.
נקודות עבורן $x_2 < x_1$ יתויגו $y = 1$, ונקודות עבורן $x_2 > x_1$ יתויגו $y = 0$.

3. השתמשו בפונקציה שכתבתם בסעיף 1, ומצאו את הסיווג שקובע הפרספטרון לנקודות שהגרלתם בסעיף 2.
הציגו בגרף את המישור של $x_1, x_2$ ואת 1000 הנקודות שהגרלתן, כאשר הנקודות המתויגות לפי הפרספטרון ב־ 1 יוצגו בכחול והנקודות המתויגות לפי הפרספטרון ב־ 0 יוצגו באדום. ציירו על המישור את וקטור המשקולות שמצאה הפונקציה. חשבו (באופן אנליטי) את הישר המפריד מתוך הוקטור שהתקבל וציירו גם אותו על המישור.
<u>**הערה**</u>: אתחלו את האלגוריתם עם תנאי ההתחלה $\bar{w} = [1, 1]$.

4. נרצה להעריך את השגיאה בין הוקטור $\bar{w}$ אליו התכנס האלגוריתם על סמך הדוגמאות שראה, לבין הפתרון האופטימלי $\bar{w}^*$. (מהו הפתרון האופטימלי במקרה שלנו?)
נגדיר שגיאה זו כערך המוחלט של הזווית (במעלות) בין $\bar{w}$ ל־$\bar{w}^*$.
נרצה לבדוק כיצד השגיאה הממוצעת משתנה כפונקציה של מספר הדוגמאות מהן האלגוריתם למד. לצורך כך, עבור כל ערך של $P = 5, 20, 30, 50, 100, 150, 200, 500$ הריצו $M = 100$ סימולציות (**בכל סימולציה הגרילו את הנקודות מחדש**) וחשבו את השגיאה הממוצעת על פני $M$ הסימולציות. הציגו גרף של השגיאה הממוצעת כפונקציה של $P$, והסבירו את התוצאה.

<u>תזכורת</u>: קוסינוס הזווית בין שני וקטורים $\bar{u}, \bar{v}$ הוא $\frac{\bar{u}^T \bar{v}}{\|\bar{u}\| \cdot \|\bar{v}\|}$.

במטלב, הפונקציה acosd היא הפונקציה ההופכית לקוסינוס שמחזירה ערכים במעלות. בפייתון, ניתן להשתמש בפונקציות np.arccos ו־ np.rad2deg.

## אלגוריתם הלמידה של הפרספטרון - תזכורת

**קלט**: נקודות $\bar{x}^1,\dots,\bar{x}^P$ והסיווג לכל נקודה $y^1,\dots,y^P$ ($y^\mu \in \{0,1\}$)
**אתחול**: אתחלו את $\bar{w}$ לוקטור כלשהו
**איטרציות**:

1. עברו על כל הנקודות בסדר מסוים קבוע
2. לכל נקודה $\bar{x}^\mu$, אם המסווג טועה לגביה, עדכנו את וקטור המשקולות באופן הבא:

$$\bar{w}\left(k+1\right)=\bar{w}\left(k\right)+\left(2y^\mu-1\right)\bar{x}^\mu$$

3. עצרו כאשר כל הנקודות מסווגות נכונה, כלומר לא היה שינוי בערך של $\bar{w}$ בכל האיטרציה האחרונה

הערה: במידה ומאתחלים את $\bar{w}$ להיות וקטור האפס, הדוגמה הראשונה שמציגים תמיד תסווג כטעות.