Machine learning

Artificial Neural Networks

Exercise VIII

פיתוח:
ד"ר יהונתן שלר
משה פרידמן

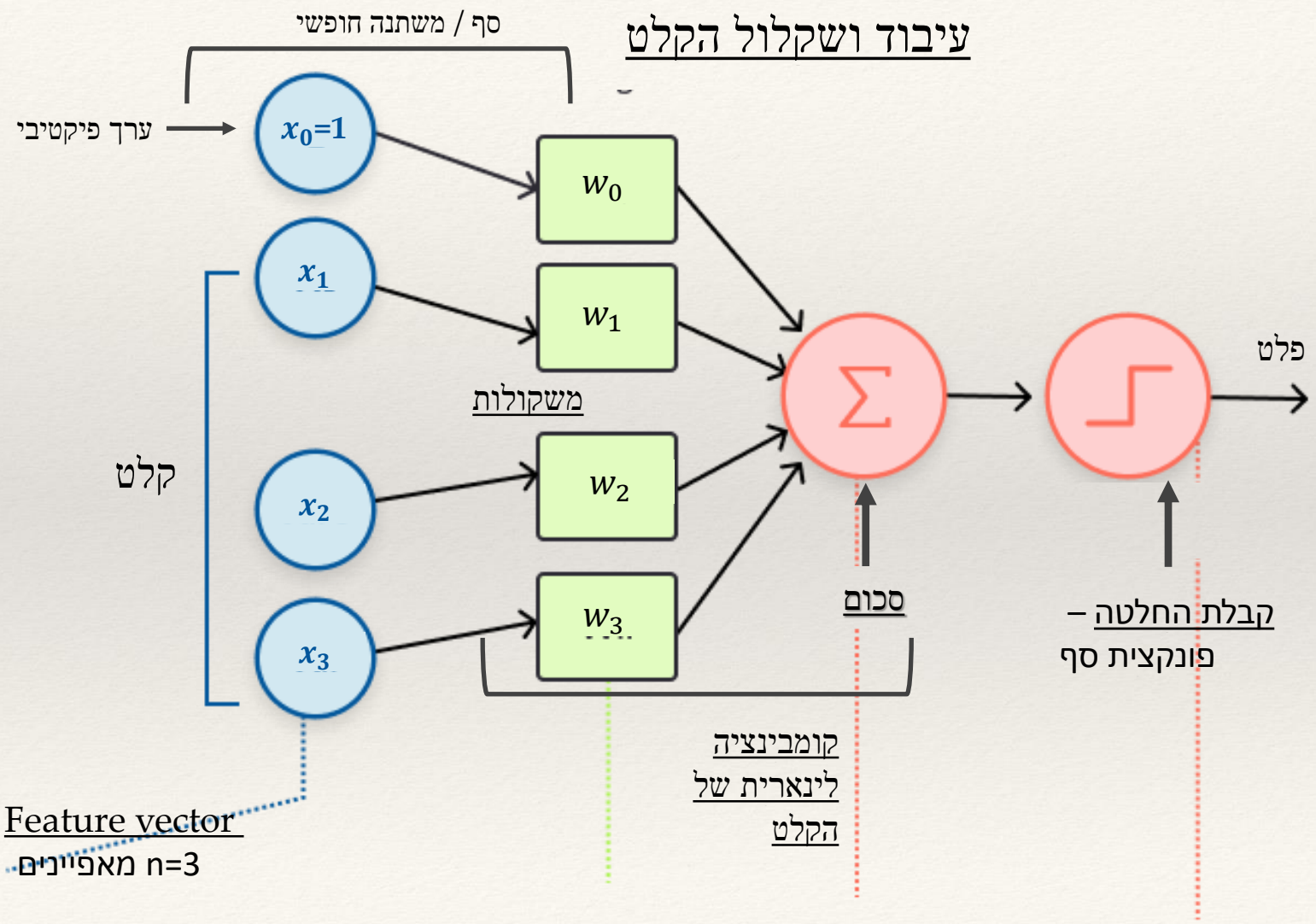
שאלה 1 - Perceptron – הנוירון המלאכותי – פונקצית ההפעלה signum

נתון פרצפטרון עם המשקולות הבאות:

w_0=-1
w_1= 1
w_2= 1
w_3= 2

שאלה - מה יהיה הסיווג של ה- feature vector הבא:

x_1=-1
x_2=-1
x_3= 2



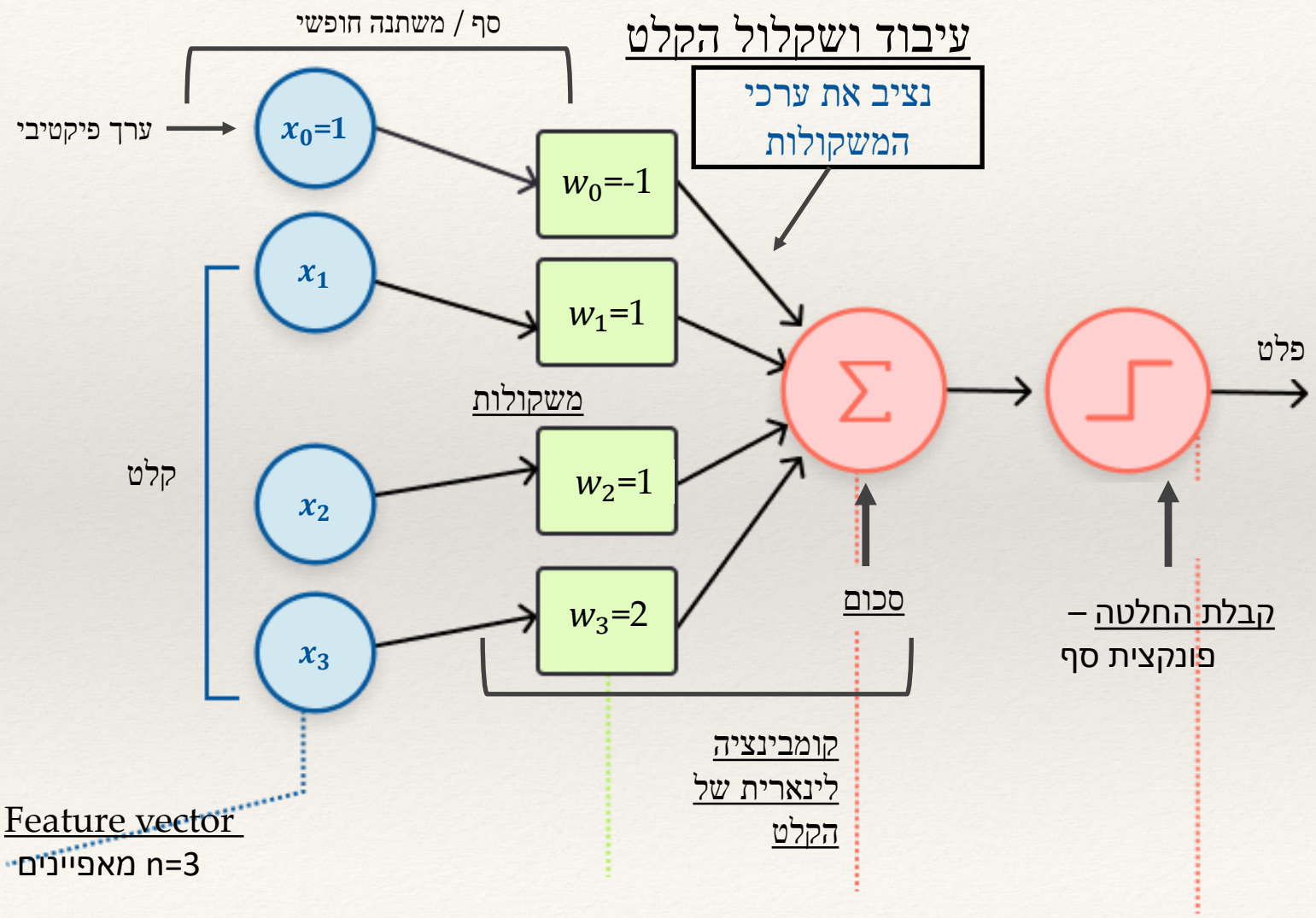
שאלה 1 - Perceptron – הנוירון המלאכותי – פונקצית ההפעלה signum

נתון פרצפטרון עם המשקולות הבאות:

w_0=-1
w_1= 1
w_2= 1
w_3= 2

שאלה - מה יהיה הסיווג של ה- feature vector הבא:

x_1=-1
x_2=-1
x_3= 2



שאלה 1 - Perceptron – הנוירון המלאכותי – פונקצית ההפעלה signum

נתון פרצפטרון עם המשקולות הבאות:

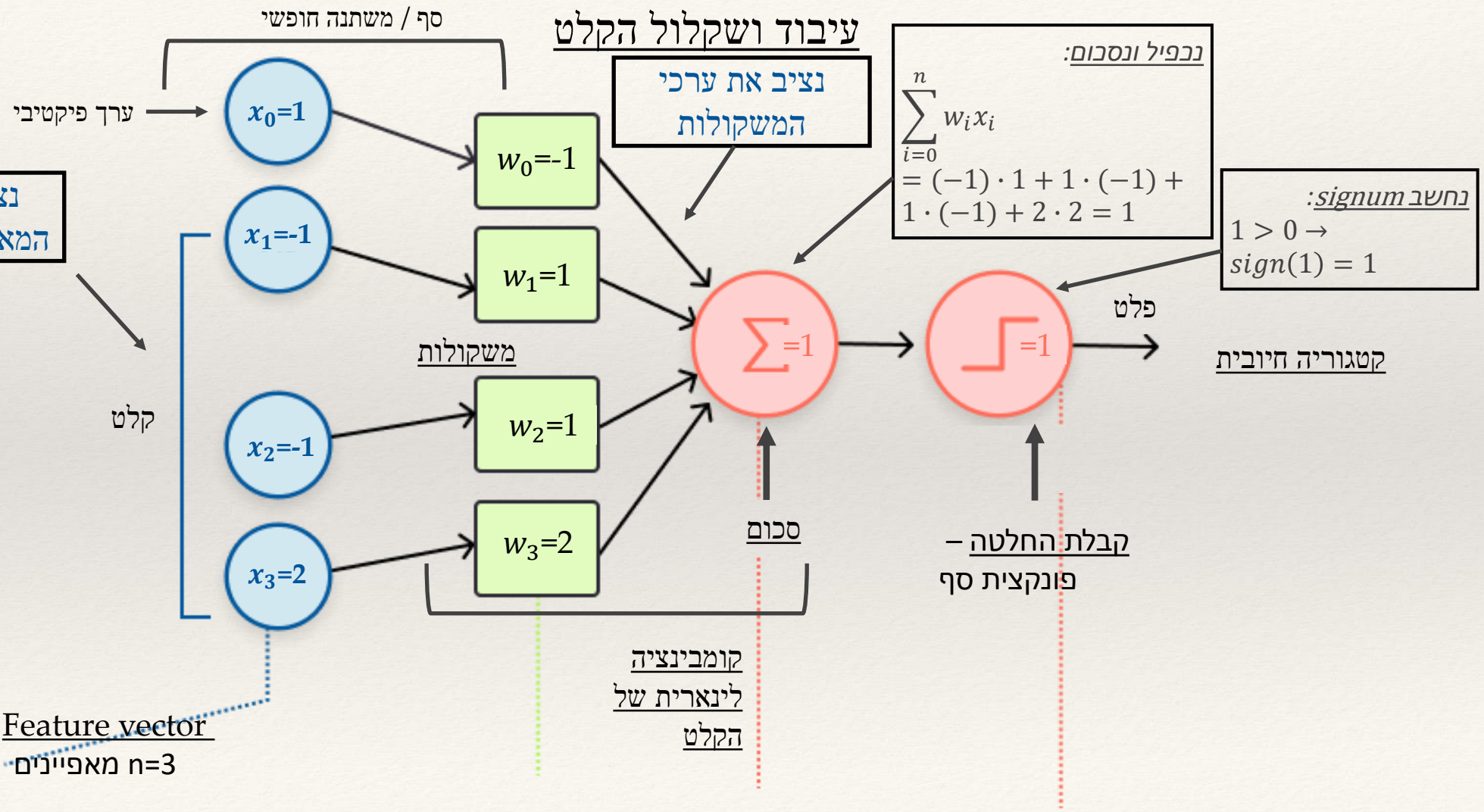
w_0=-1
w_1= 1
w_2= 1
w_3= 2

נציב את ערכי המאפיינים (הקלט)

שאלה - מה יהיה הסיווג של ה- feature vector הבא:

x_1=-1
x_2=-1
x_3= 2

תשובה:
קטגוריה חיובית



שאלה 2 - Perceptron – הנוירון המלאכותי – פונקצית ההפעלה signum

נתון פרצפטרון עם המשקולות הבאות:

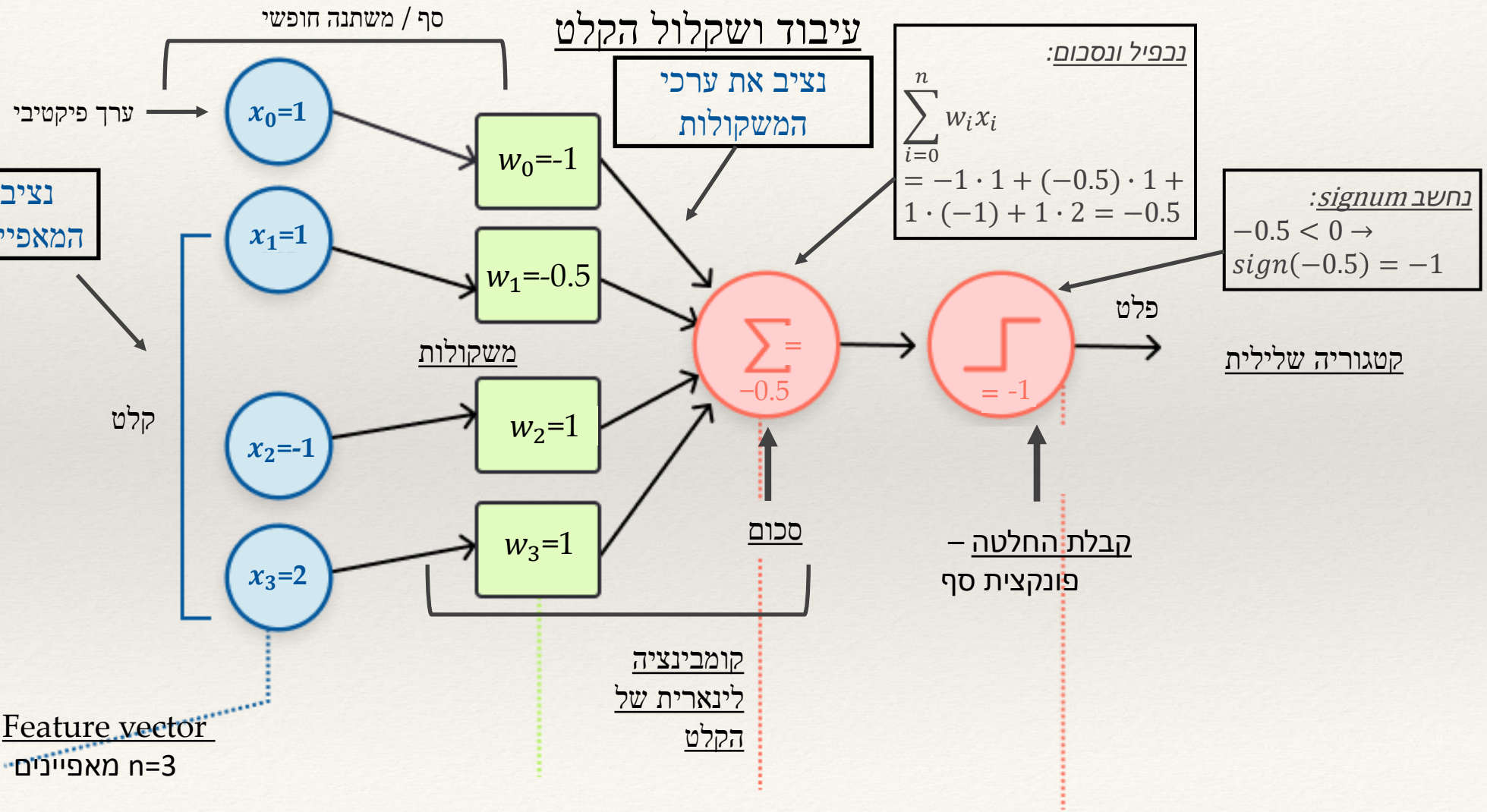
w_0=-1
w_1= -0.5
w_2= 1
w_3= 1

נציב את ערכי המאפיינים (הקלט)

שאלה - מה יהיה הסיווג של ה- feature vector הבא:

x_1= 1
x_2=-1
x_3= 2

תשובה:
קטגוריה חיובית

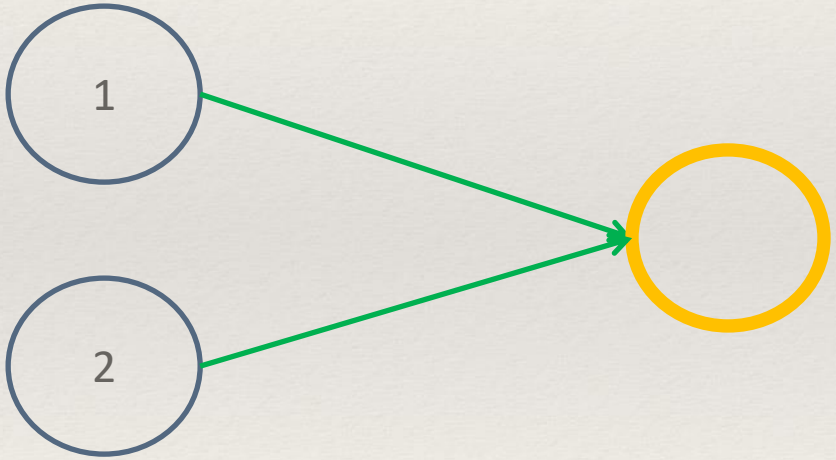


- Feature vector
n=3 מאפיינים

שאלה 3 – יחידה בודדת – Perceptron עם sign

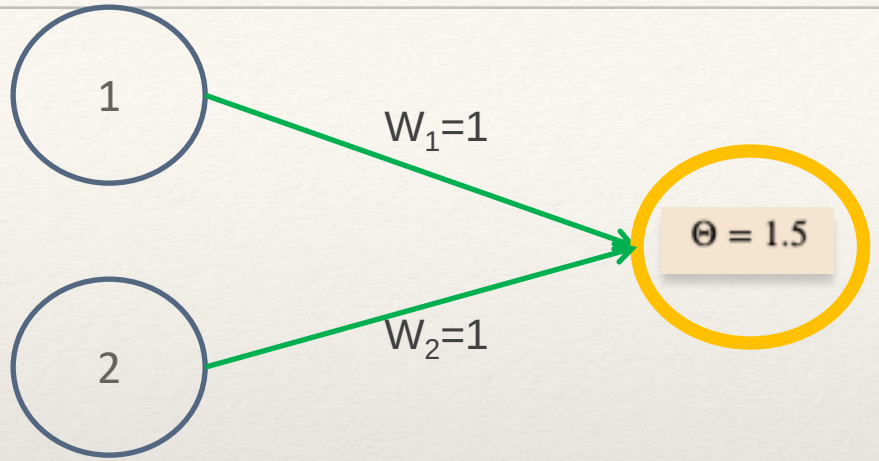
- בנו רשת במבנה הר"מ שמחשבת AND לוגי של 2 משתנים לפי טבלת האמת הרצ"ב.
- מהם משקלות הרשת (w0,w1,w2)?

AND		
in_1	in_2	out_1
0	0	0
0	1	0
1	0	0
1	1	1



שאלה 3 – יחידה בודדת – Perceptron עם sign – פתרון (אחד מאינסוף פתרונות)

AND		
in_1	in_2	out_1
0	0	0
0	1	0
1	0	0
1	1	1



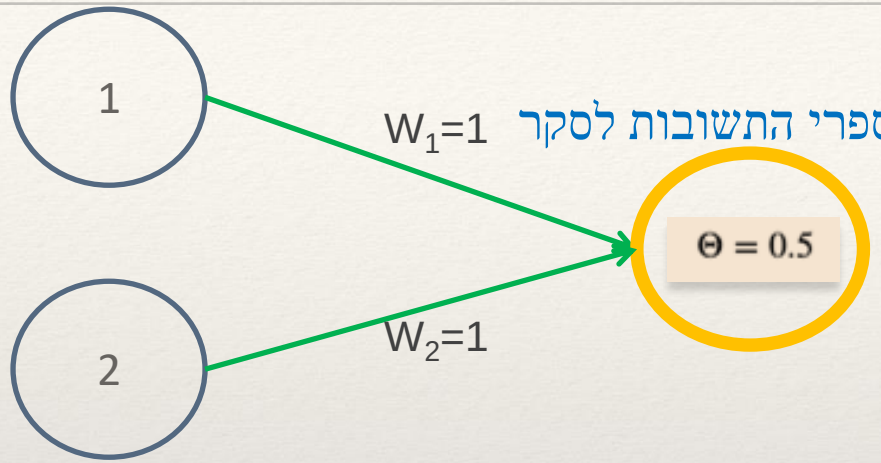
$$out_j = \text{sgn}\left(\sum_{i=1}^{n} in_i w_{ij} - \theta\right)$$

j = 1 , n = 2

$$out(0,0) = \text{sgn}(0 \times 1 + 0 \times 1 - 1.5) = \text{sgn}(-1.5) = 0$$
$$out(0,1) = \text{sgn}(0 \times 1 + 1 \times 1 - 1.5) = \text{sgn}(-0.5) = 0$$
$$out(1,0) = \text{sgn}(1 \times 1 + 0 \times 1 - 1.5) = \text{sgn}(-0.5) = 0$$
$$out(1,1) = \text{sgn}(1 \times 1 + 1 \times 1 - 1.5) = \text{sgn}(0.5) = 1$$

שאלת סקר - שאלה 4 – יחידה בודדת – Perceptron עם sign

???		
in_1	in_2	out_1
0	0	?
0	1	?
1	0	?
1	1	?



מספרי התשובות לסקר

	תשובות אפשריות		
4	3	2	1
0	0	0	1
0	1	1	1
0	1	0	1
0	1	1	1

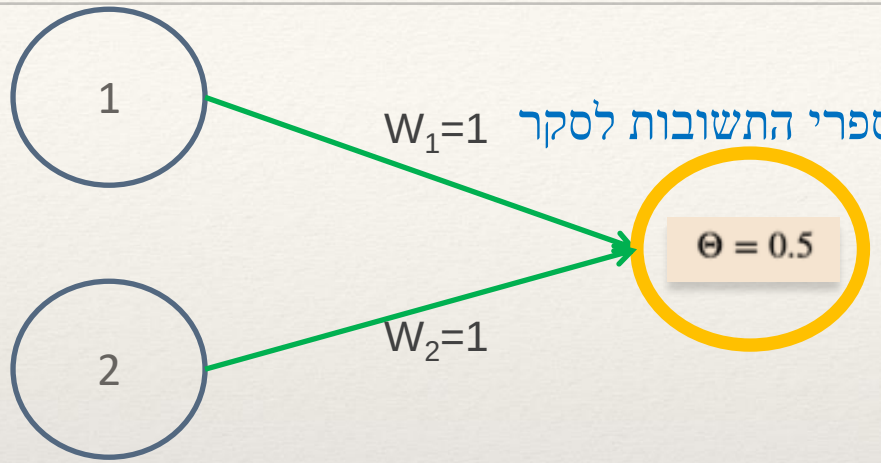
out for 1st example
out for 2nd example
out for 3rd example
out for 4th example

שאלת סקר: מהם הסיווגים שיתקבלו מה-perceptron המצויר?

- שימו לב – מספרי התשובה מופיעים בכותרות העמודות

שאלת סקר - שאלה 4 – יחידה בודדת – Perceptron עם sign

???		
in_1	in_2	out_1
0	0	?
0	1	?
1	0	?
1	1	?



מספרי התשובות לסקר

	תשובות אפשריות		
4	3	2	1
0	0	0	1
0	1	1	1
0	1	0	1
0	1	1	1

out for 1st example

out for 2nd example

out for 3rd example

out for 4th example

Bias = −1.5			Bias = −0.5		
Input	Weighted sum	Output	Input	Weighted sum	Output
(0, 0)	−1.5	0	(0, 0)	−0.5	0
(0, 1)	−0.5	0	(0, 1)	0.5	1
(1, 0)	−0.5	0	(1, 0)	0.5	1
(1, 1)	0.5	1	(1, 1)	1.5	1

אימון – training

- תיקון המשקולות

אימון פרצפטרון – עדכון המקדמים

- כלל העדכון של המשקולות:

$$w_i \leftarrow w_i + \Delta w_i$$

$$\Delta w_i = \eta(t - o)x_i$$

Δw_i - עדכון ל-w_i (הוספה או הורדה ביחס לאיטרציה הקודמת)

η קבוע (קטן מ-1) הקובע את קצב הלמידה (למשל 0.1)

t סימון הדוגמא הנוכחית - ערך הקטגוריה האמיתי של הדוגמה

o הערך שנותן ה perceptron עבור הדוגמא הנוכחית

Gradient Descent בפרצפטרון (with sign function) – פסאודו קוד

סימונים:

η קבוע (קטן מ-1) הקובע את קצב הלמידה (למשל 0.1)

t סימון הדוגמא הנוכחית - ערך הקטגוריה האמיתי של הדוגמה

o הערך שנותן ה נוירון עבור הדוגמא הנוכחית

$\langle \vec{x}, t \rangle$ - נסמן כל דוגמה כזוג feature vector וקטגוריה

Δw_i - עדכון ל-w_i (הוספה ביחס לאיטרציה הקודמת)

אלגוריתם:

Gradient Descent (training examples, η):

- לכל משקולת w_i, אתחל ערכים התחלתיים אקראיים קטנים
- מעדכנים את הפרמטרים עד להתכנסות:
  - אתחל כל Δw_i ל-0
  - עבור כל $\langle \vec{x}, t \rangle$ בדוגמאות האימון
    - חשב את o ע"י הכנסת $\vec{x}$ כקלט ליחידת הנוירון
    - לכל משקולת w_i:
      - $\Delta w_i = \Delta w_i + \eta(t-o)x_i$
  - $w_i^{t+1} := w_i^t + \Delta w_i$

שאלה 5א – פרצפטרון – שלב האימון

נתון

נתון קבוע הלמידה: η=0.1

נתונים מקדמי המודל באמצע האימון:

$$w_0 = -5, w_1 = 1, w_2 = 1$$

כלומר המפריד הלינארי:

1*x1+1*x2-5

עבור הווקטור הבא:

$$x_0 = 1, x_1 = 1, x_2 = 1$$

הקטגוריה האמתית הנתונה – היא שלילית

שאלות: 1. מה יהיה ערך הפונקציה sign?

תשובה: $-5 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 < 0$

לכן פונקצית sign = -1

---

2. מהם ערכי o ו-t?

תשובה: o=-1, t=-1

---

3. מה יהיו ערכי המקדמים לאחר העדכון?

תשובה: המקדמים לא ישתנו, כיוון ש- o=t

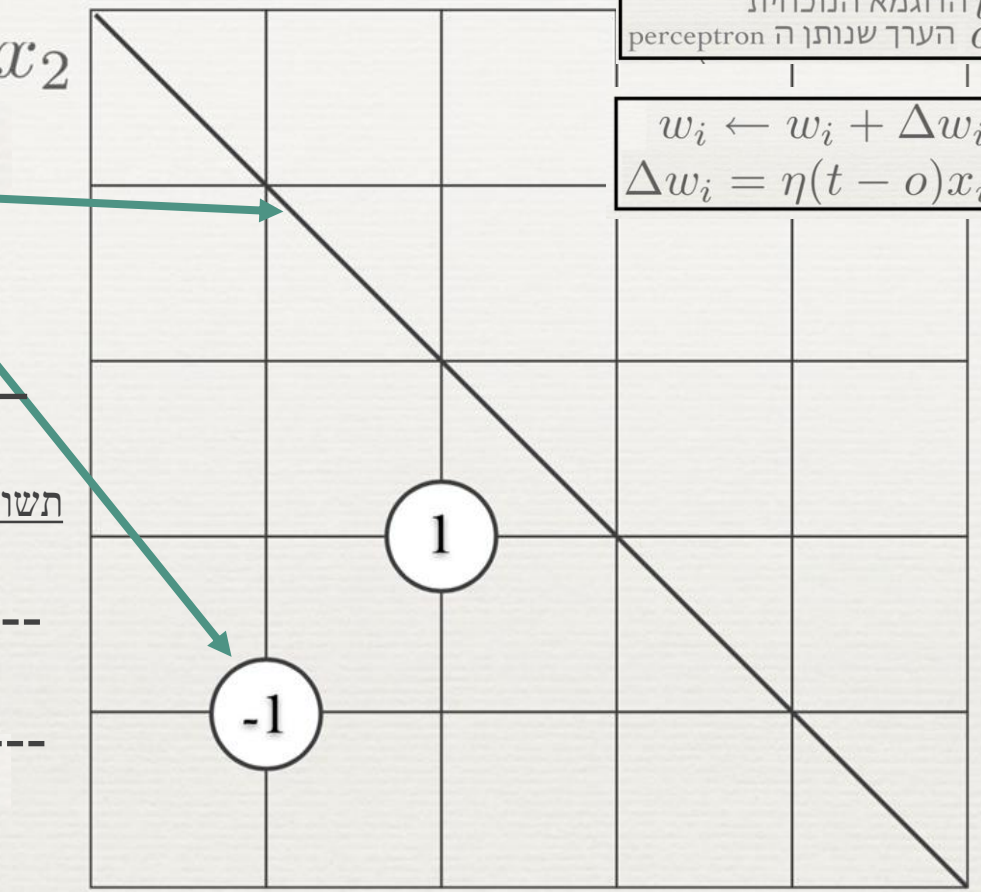
η קבוע הלמידה

t הדוגמא הנוכחית

o הערך שנותן ה perceptron

$$w_i \leftarrow w_i + \Delta w_i$$

$$\Delta w_i = \eta(t - o)x_i$$



ראשית הצירים: 0,0

שאלה 5ב – פרצפטרון – שלב האימון

נתון

נתון קבוע הלמידה: η=0.1

נתונים מקדמי המודל באמצע האימון:

$$w_0 = -5, w_1 = 1, w_2 = 1$$

כלומר המפריד הלינארי:

1*x1+1*x2-5

עבור הווקטור הבא:

$$x_0 = 1, x_1 = 2, x_2 = 2$$

הקטגוריה האמתית הנתונה – היא חיובית

שאלות: 1. מה יהיה ערך הפונקציה sign?

תשובה: -5 · 1+1·2+1·2=-1<0

לכן פונקצית sign = -1

2. מהם ערכי o ו-t? תשובה: o=-1, t=1

3. מה יהיו ערכי המקדמים לאחר העדכון?

תשובה: o≠t, לכן נעדכן את המקדמים.

$\Delta w_i = \eta\cdot(t-o)\cdot x_i$ $\qquad w_i := w_i + \Delta w_i$

$\Delta w_0 = 0.1\cdot(1-(-1))\cdot 1 = 0.2 \rightarrow w_0 := -5+0.2 = -4.8$

$\Delta w_1 = 0.1\cdot(1-(-1))\cdot 2 = 0.4 \rightarrow w_1 := 1 + 0.4 = 1.4$

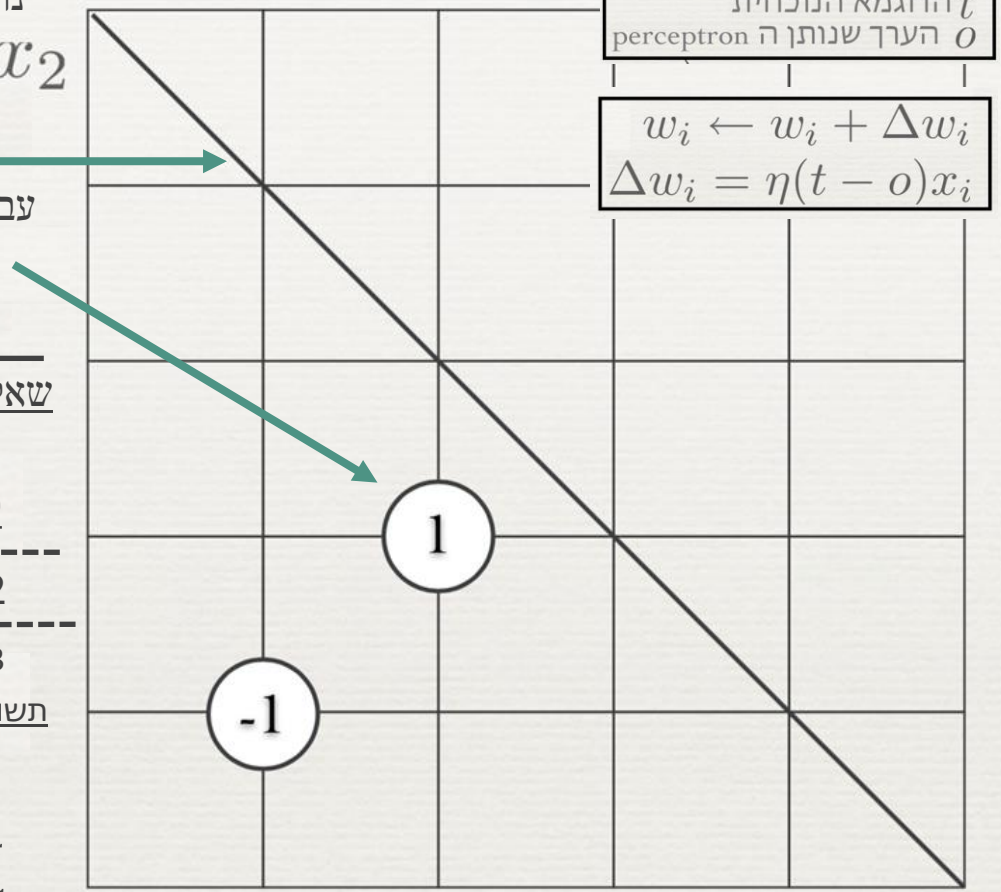
$\Delta w_2 = 0.1\cdot(1-(-1))\cdot 2 = 0.4 \rightarrow w_2 := 1 + 0.4 = 1.4$

η קבוע הלמידה

t הדוגמא הנוכחית

o הערך שנותן ה perceptron

$$w_i \leftarrow w_i + \Delta w_i$$
$$\Delta w_i = \eta(t-o)x_i$$



ראשית הצירים: 0,0

x_1

שאלה 5ג – פרצפטרון – שלב האימון

נתון

נתון קבוע הלמידה: η=0.1

נתונים מקדמי המודל באמצע האימון:

$w_0 = -4.8, w_1 = 1.4, w_2 = 1.4$

כלומר המפריד הלינארי:

1.4*x1+1.4*x2-4.8

עבור הוקטור הבא:

$x_0 = 1, x_1 = 2, x_2 = 2$

הקטגוריה האמיתית הנתונה – היא חיובית

שאלות: 1. מה יהיה ערך הפונקציה sign?

תשובה: $-4.8 \cdot 1+1.4 \cdot 2+1.4 \cdot 2=0.8>0$

לכן פונקצית sign = 1

---

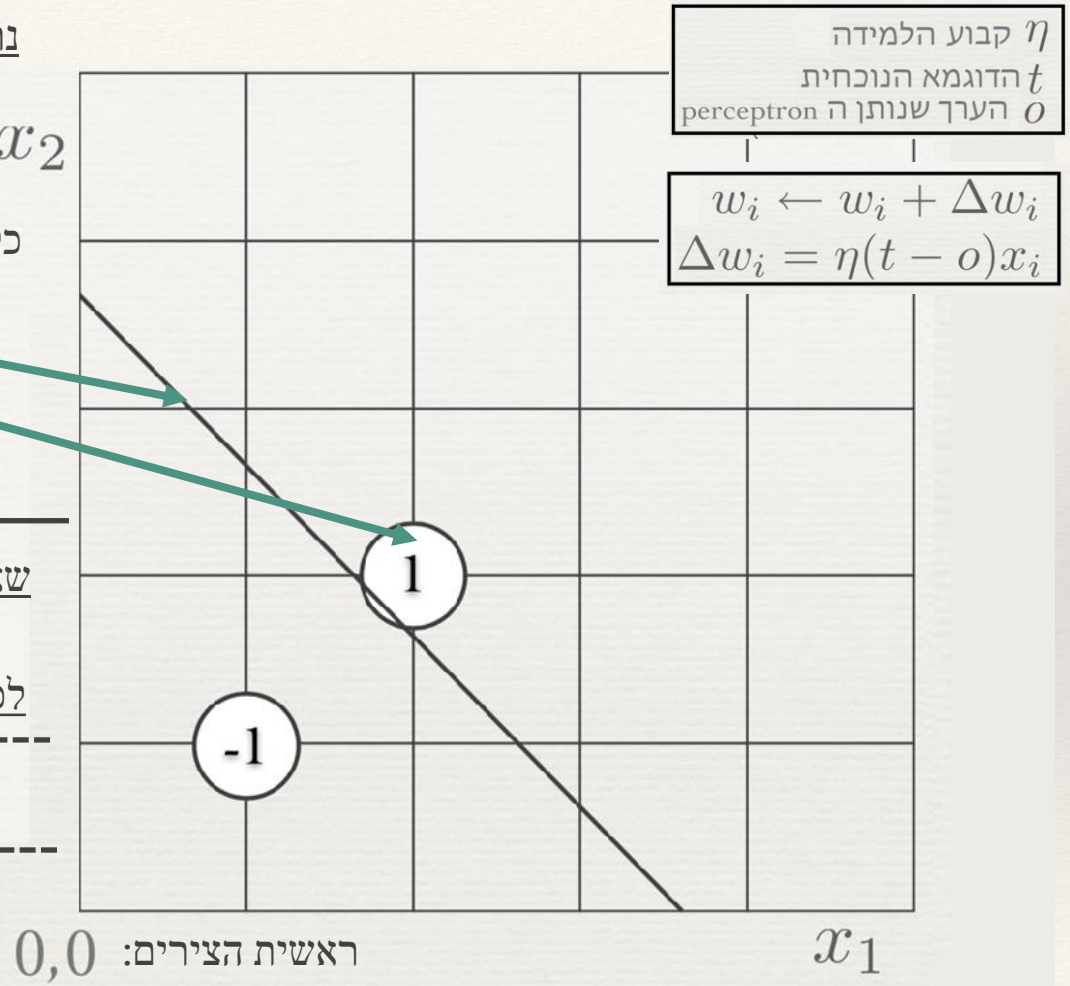
2. מהם ערכי o ו-t?

תשובה: o=1, t=1

---

3. מה יהיו ערכי המקדמים לאחר העדכון?

תשובה: המקדמים לא ישתנו, כיוון ש- o=t



ראשית הצירים:

η קבוע הלמידה
t הדוגמא הנוכחית
o הערך שנותן ה perceptron

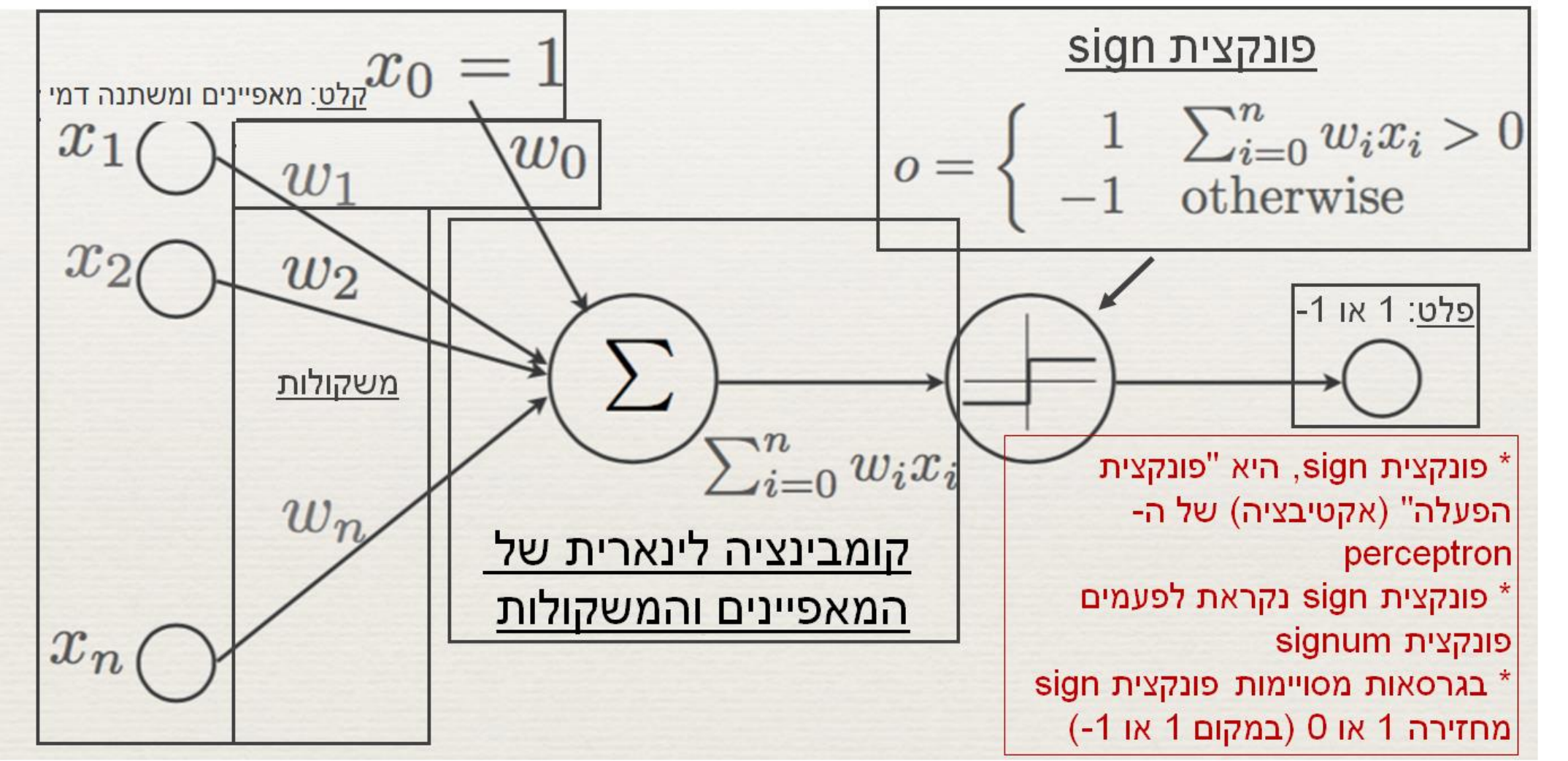
$$w_i \leftarrow w_i + \Delta w_i$$
$$\Delta w_i = \eta(t-o)x_i$$

רשתות נוירונים

טופולוגיות של רשתות נוירונים וסיווג (feed forward)

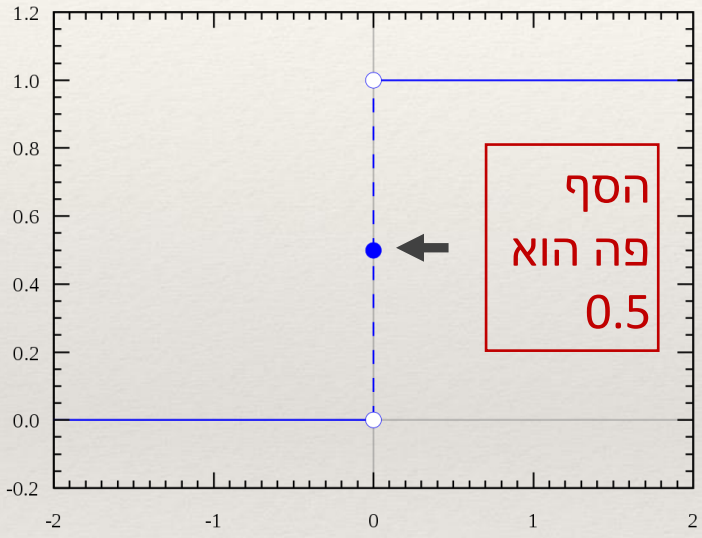
- יחידה בודדת לינארית
- יחידה בודדת – perceptron
- יחידה בודדת – נוירון עם sigmoid
- רשת רב שכבתית

Perceptron – מודל א'
הנוירון המלאכותי עם פונקצית הפעלה sign

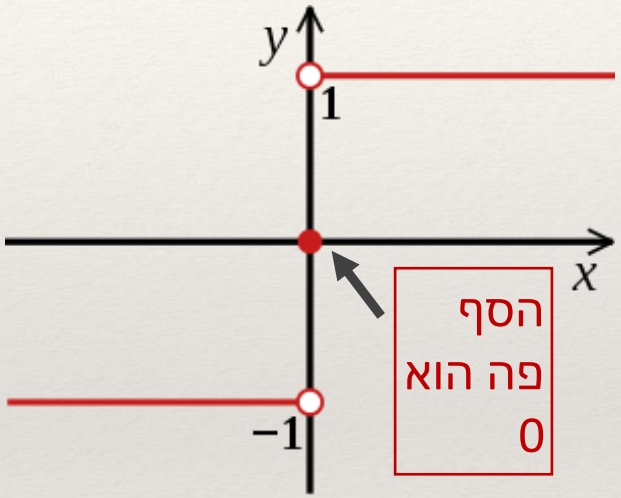


Perceptron – מודל א' – פונקצית הפעלה

<u>פונקצית הסף</u> – פונקצית ההפעלה היא פונקצית מדרגה ובעצם ראינו 2 כאלה:

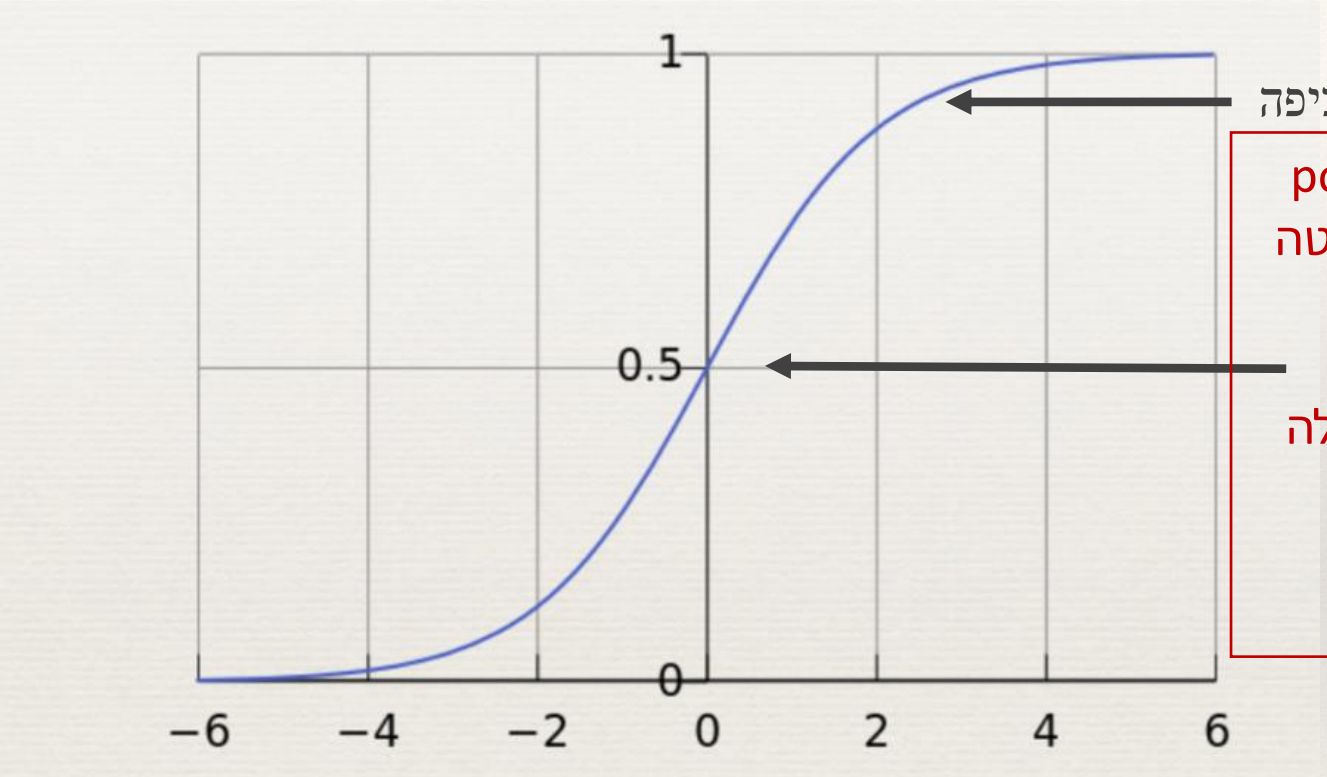


Heaviside step function
(פונקצית מדרגה)



Signum function
(פונקצית הסימן – sign)

יחידה בודדת – נוירון עם "פונקצית ההפעלה" - sigmoid *



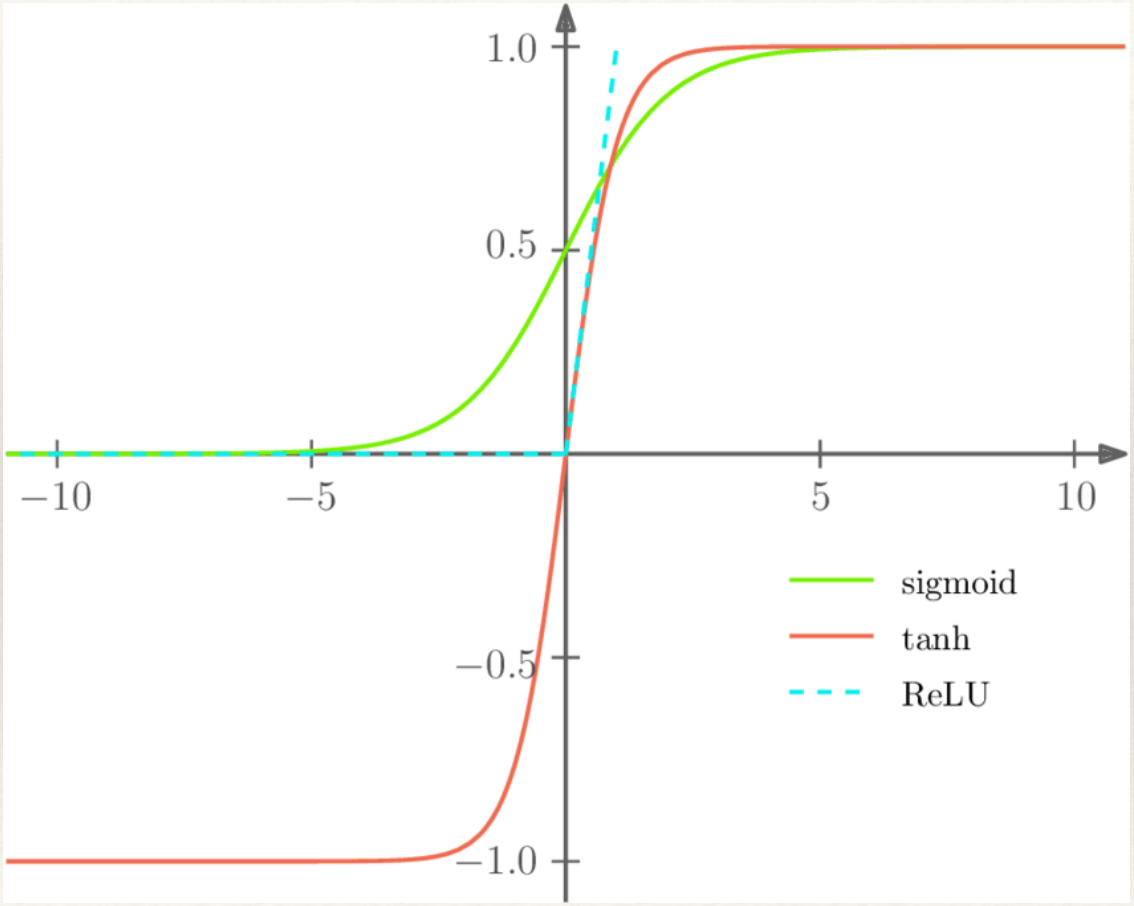
פונקצית ה-sigmoid רציפה

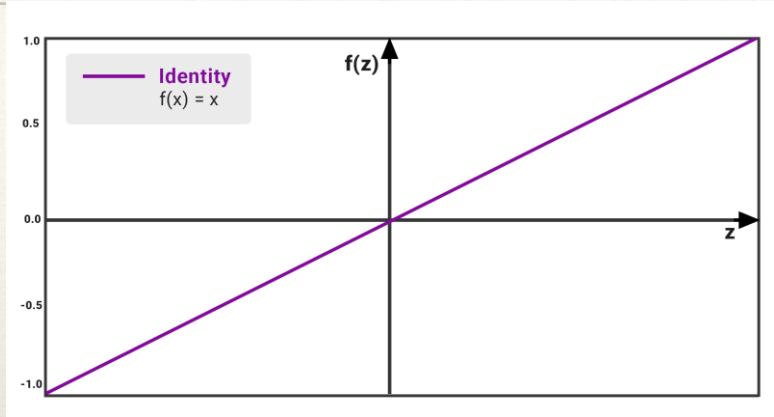
* מוסיפים ב-post processing שלב בדיקת סף, לקבלת החלטה על ערך הקטגוריה
* הסף פה הוא 0.5 (ולא 0)
- אם תוצאת הפונקציה גדולה מ-0.5,
- נסווג בקטגוריה החיובית, אחרת, בקטגוריה שלילית

$\sigma(x)$ is the sigmoid function

$$\frac{1}{1+e^{-x}}$$

פונקציות אקטיבציה נוספות





sigmoid

$s(z) = \frac{1}{1+e^{-z}}$

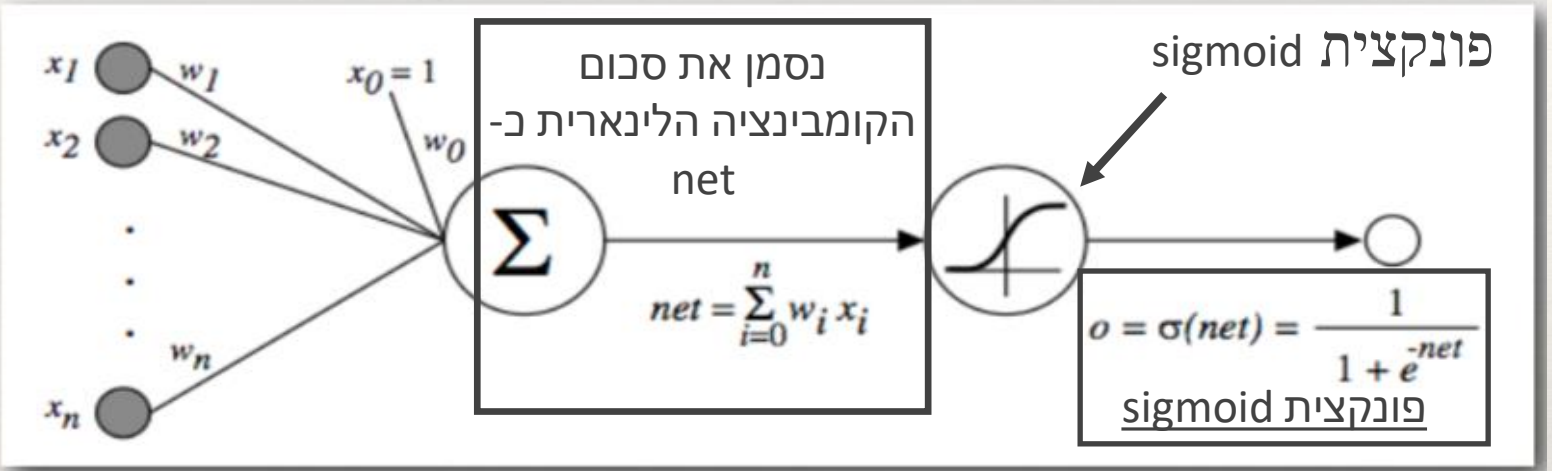
tanh

$\tanh(z) = \frac{e^{2z}-1}{e^{2z}+1}$

ReLU (Rectified Linear Unit)

$R(z) = \max(0,z)$

יחידה בודדת – נוירון עם "פונקצית ההפעלה" - sigmoid *



$\sigma(x)$ is the sigmoid function

$$\frac{1}{1 + e^{-x}}$$

פונקצית ה-sigmoid
שימו לב שפה ה-x משמש כסכום
הקומבינציה הלינארית

שאלה 6 - דוגמה לקריאת נוירון בודד עם sigmoid

$$S(x) = \frac{1}{1+e^{-x}}$$

נתונה יחידת הנוירון הבאה, כאשר מישור ההפרדה מיוצג ע"י:
$0.2 \cdot x_1 - 0.5 \cdot x_2 + 0.4 \cdot x_3 - 2 = 0$
בצעו סימולציה לקבלת הפלט של הנוירון, תוך שימוש בגרף של נוירון ה-sigmoid, עבור הוקטור: (5,4,10)

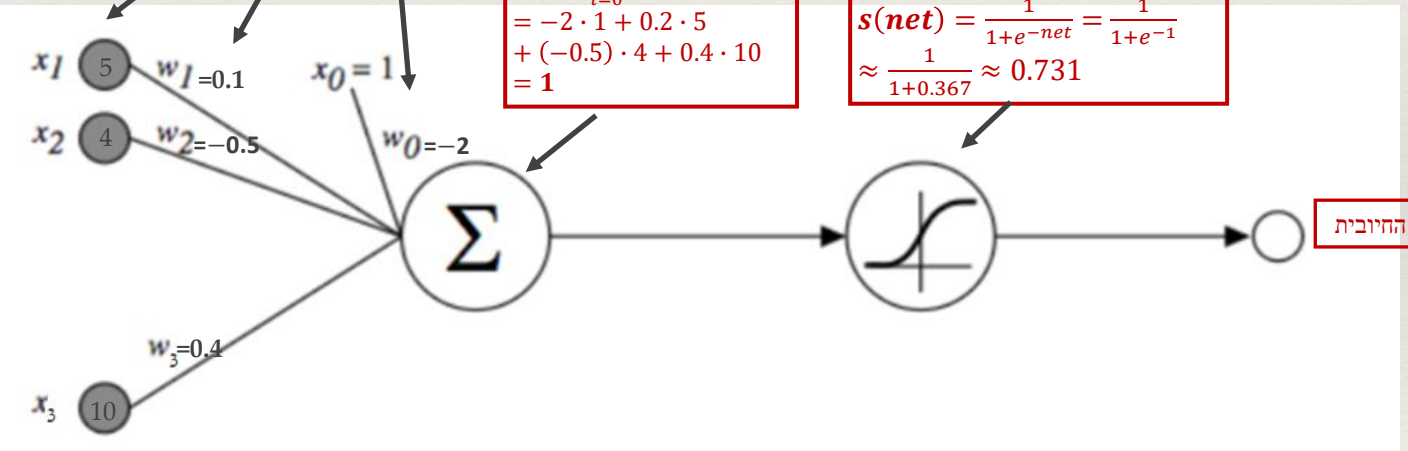
שלב א' – נציב את ערכי המשקולות וערכי ה-feature vector

שלב ב' - נכפיל ונסכום:

$$\boldsymbol{net} = \sum_{i=0}^{3} w_i x_i$$
$$= -2 \cdot 1 + 0.2 \cdot 5 + (-0.5) \cdot 4 + 0.4 \cdot 10 = \mathbf{1}$$

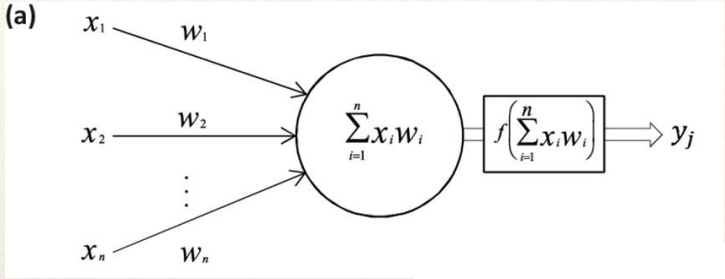
שלב ג' – נחשב את ערך הסיגמויד:

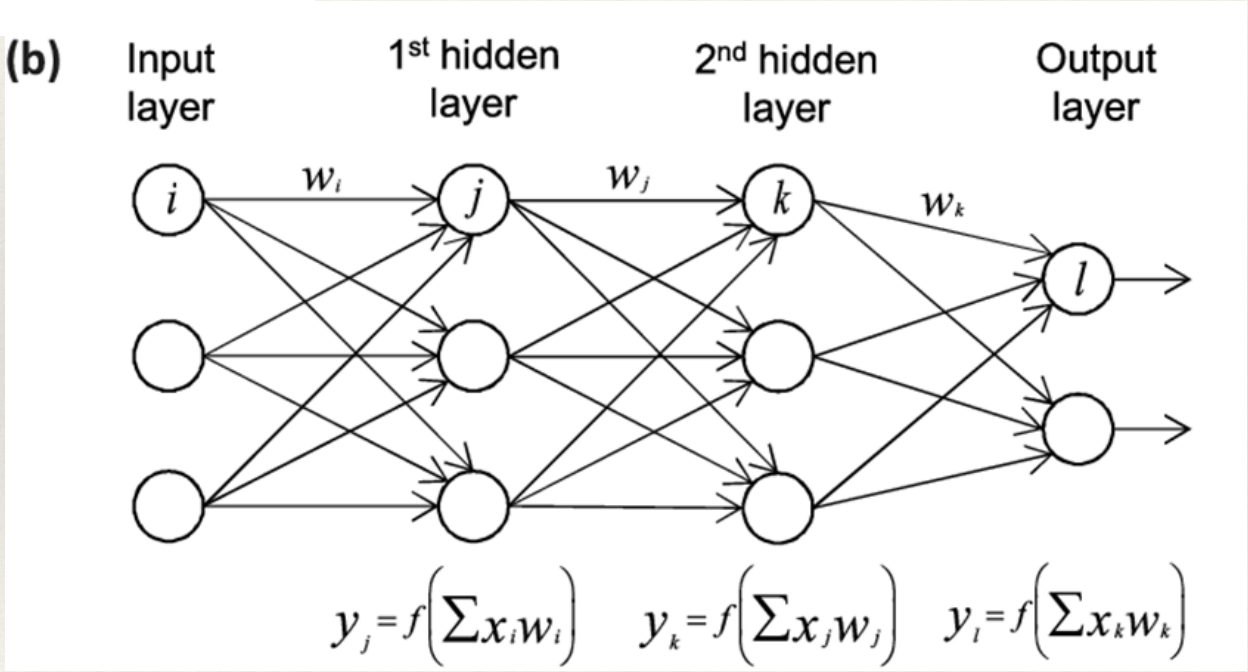
$$\boldsymbol{s(net)} = \frac{1}{1+e^{-net}} = \frac{1}{1+e^{-1}} \approx \frac{1}{1+0.367} \approx 0.731$$



הסיווג= המחלקה החיובית

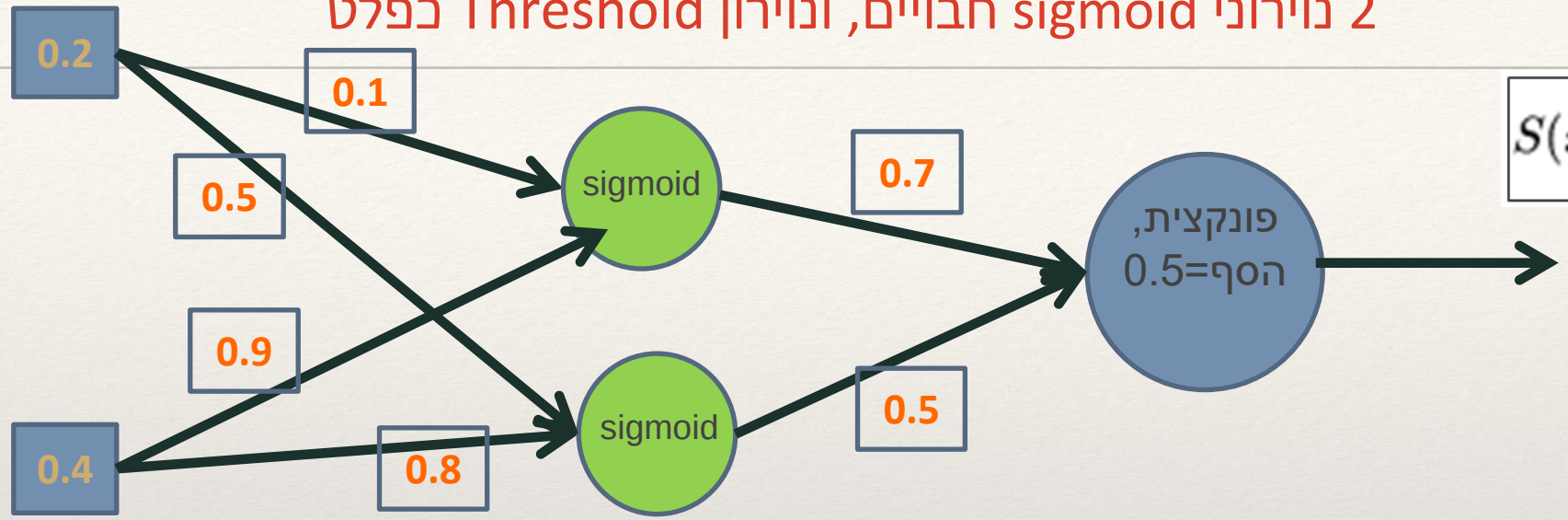
רשת רב שכבתית – בהשוואה ליחידה בודדת





שאלה 7 - רשת נוירונים רב שכבתית

2 נוירוני sigmoid חבויים, ונוירון Threshold בפלט



$$S(x) = \frac{1}{1+e^{-x}}$$

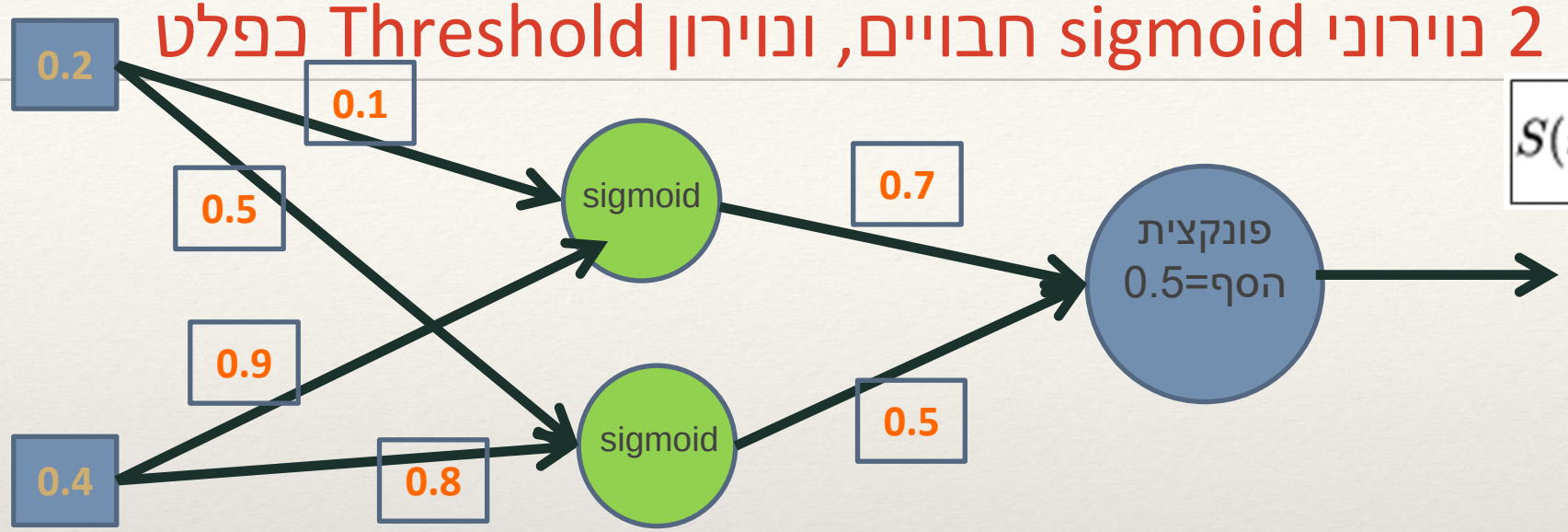
חשב את הפלט של הרשת: סקר.

<u>תשובות אפשריות</u>:

1. 1
2. 0

שאלה 7 - רשת נוירונים רב שכבתית

2 נוירוני sigmoid חבויים, ונוירון Threshold כפלט



$$S(x) = \frac{1}{1+e^{-x}}$$

Hidden 1 (0.2 x 0.1) + (0.4 x 0.9) = 0.38 $output = \frac{1}{1+e^{-0.38}} = 0.594$

Hidden 2 (0.2 x 0.5) + (0.4 x 0.8) = 0.42 $output = \frac{1}{1+e^{-0.42}} = 0.603$

Output (0.594 x 0.7) + (0.603 x 0.5) = 0.717 > 0.5 → output = 1

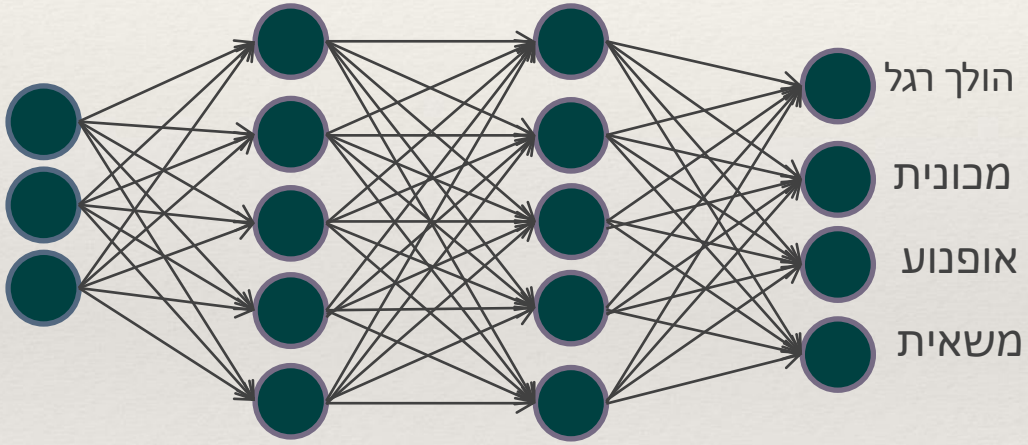
רשת רב שכבתית – מוטיבציה

Pedestrian

Car

Motorcycle

Truck

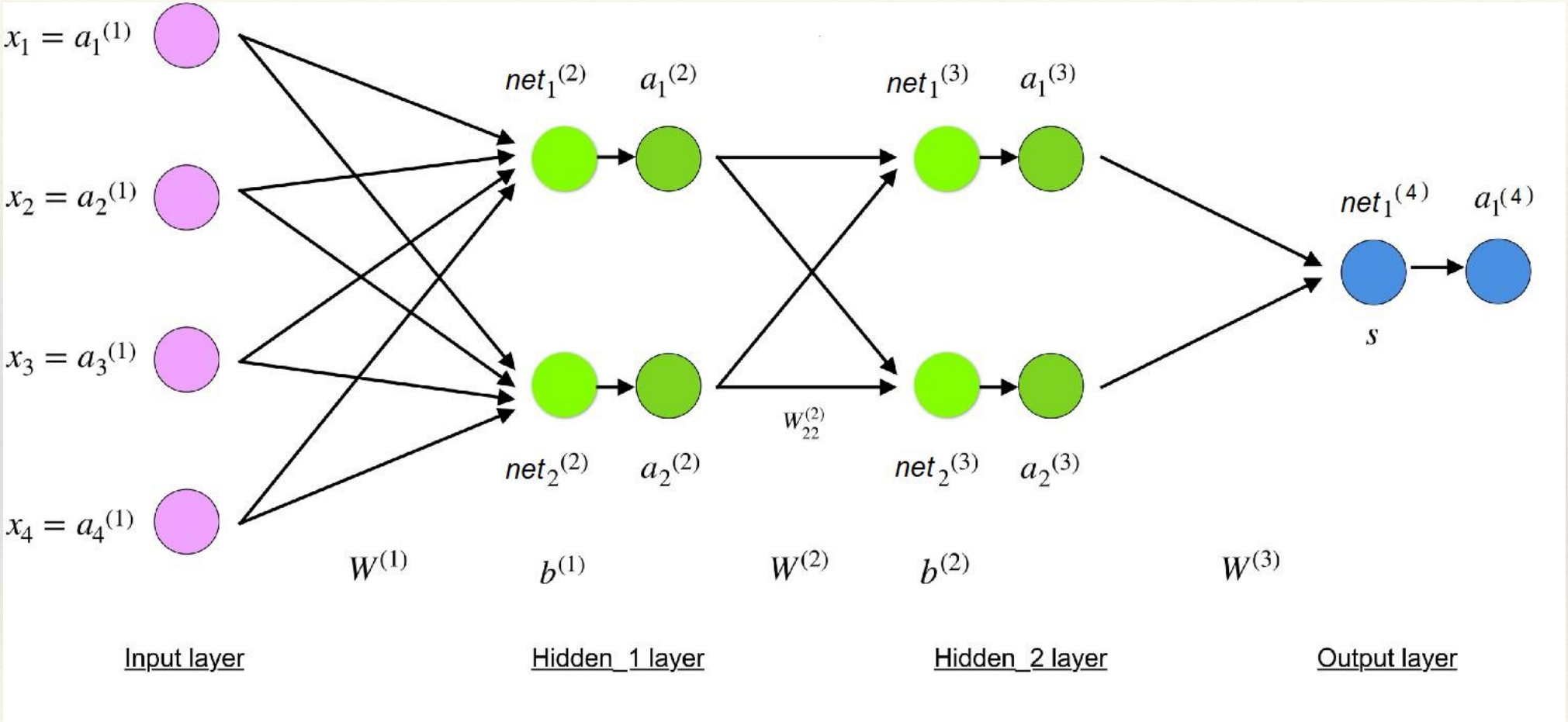


$$h_w(x) \in \mathbb{R}^4$$

$$h_w(x) \approx \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad h_w(x) \approx \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad h_w(x) \approx \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

הולך רגל מכונית אופנוע

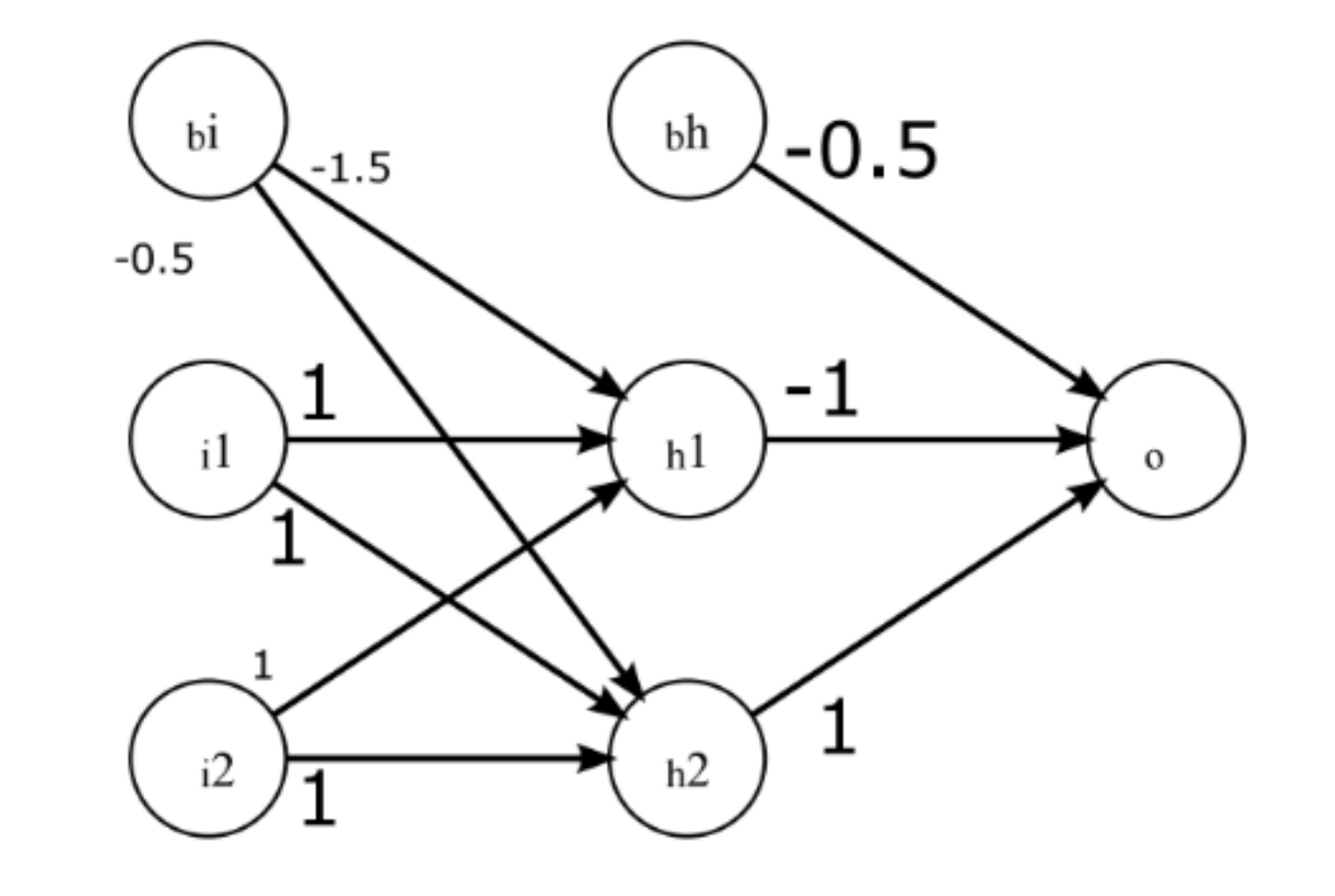
רשת רב שכבתית – עם 2 שכבות חבויות



net - סכום הקומבינציה הלינארית

a – פונקצית ההפעלה (sign או sigmoid, או פונקצית הזהות, עבור שכבת הקלט)

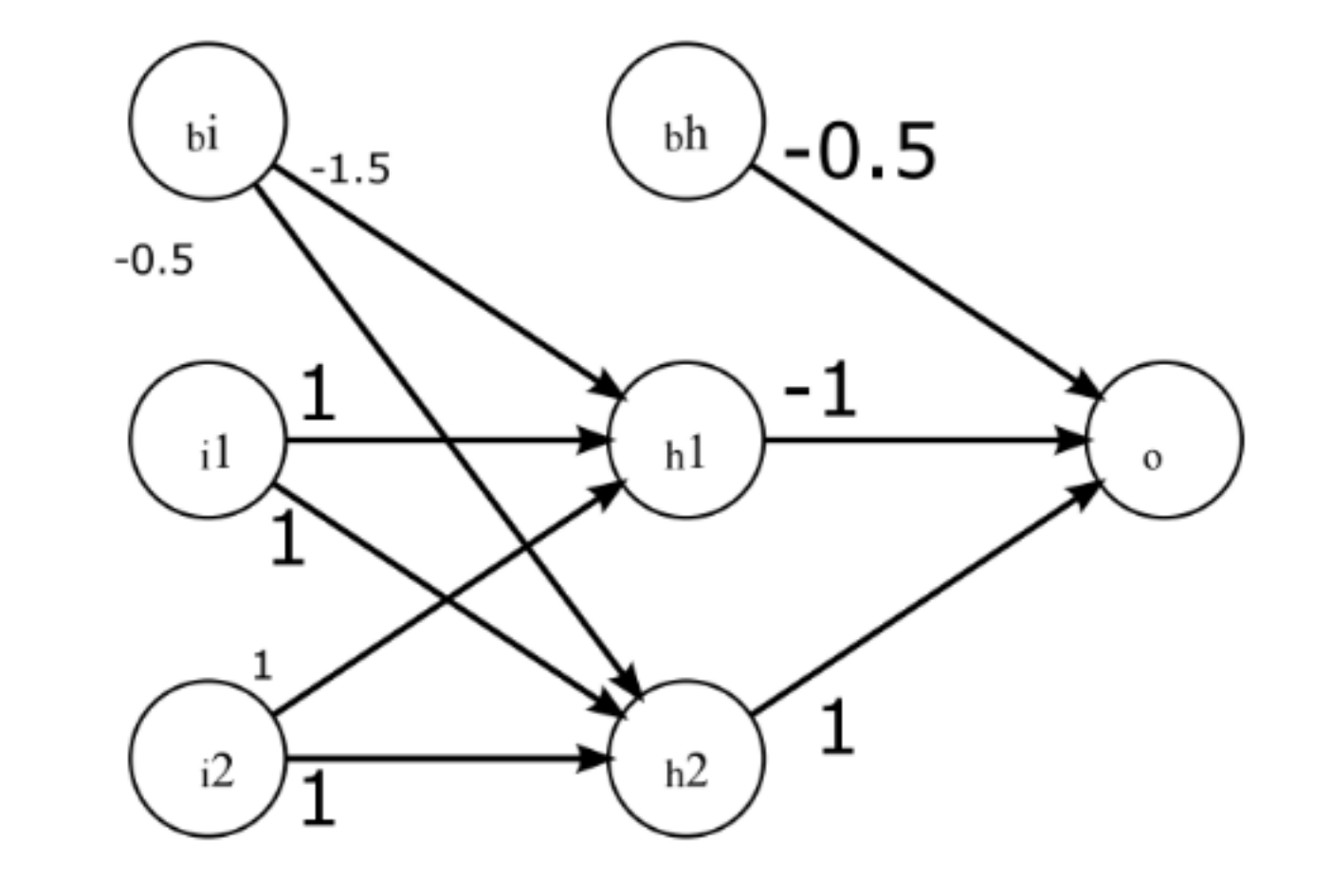
שאלה 8 - Feed forward (predict) - רשת רב שכבתית לינארית



מה מחשב
הפרצפטרון הבא?
שים לב, שכאן יש
פונק' SIGN גם
בשכבה הפנימית

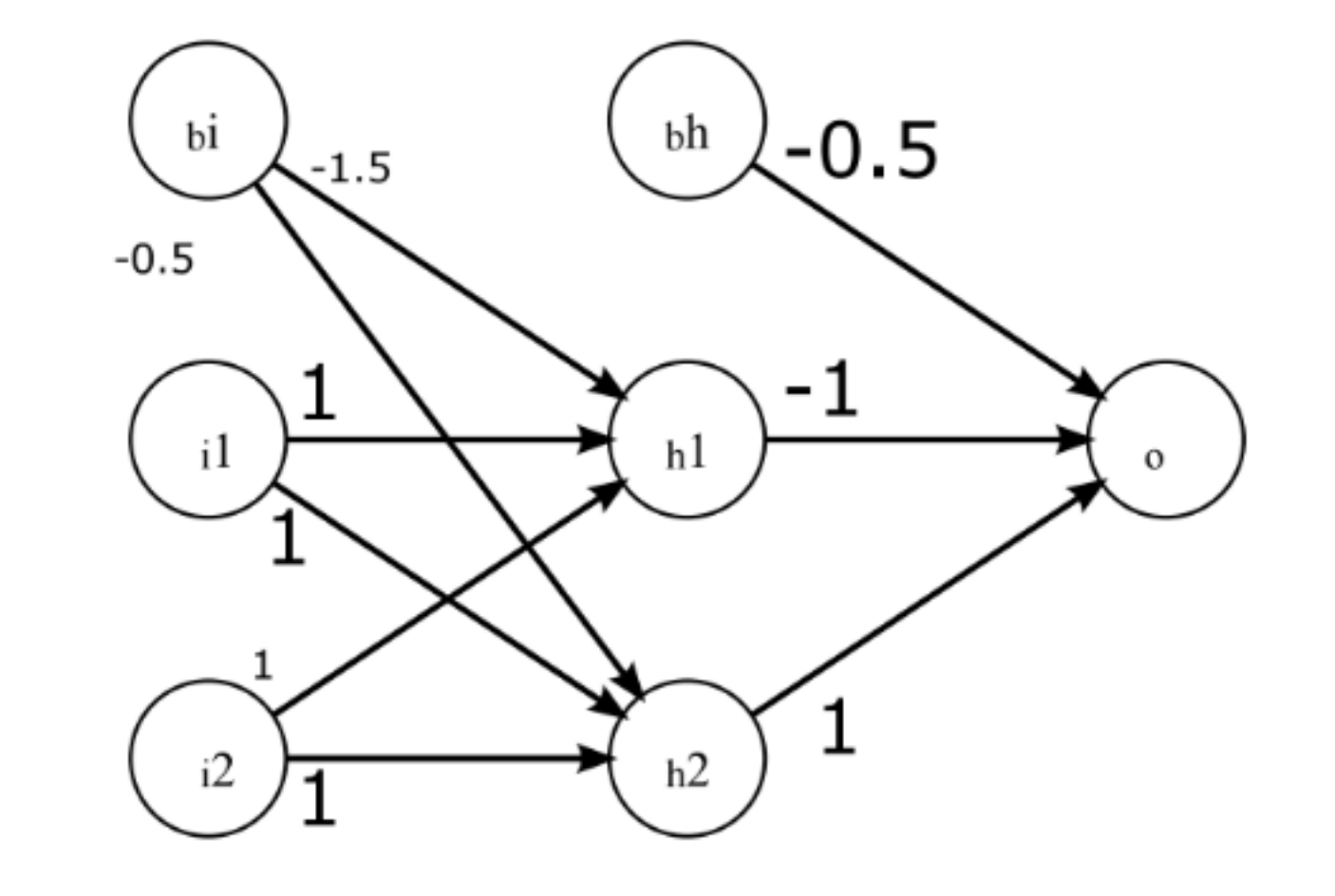
i1	i2
0	0
0	1
1	0
1	1

שאלה 8 - Feed forward (predict) - רשת רב שכבתית לינארית – פתרון



i1	i2	h1	h2	XOR
0	0	-1.5 => -1	-0.5 => -1	-.5 =>-1
0	1	-0.5 => -1	0.5 => 1	1.5 **=> 1**
1	0	-0.5 => -1	0.5 => 1	1.5 **=> 1**
1	1	.5 => 1	1.5 = > 1	-.5 =>-1

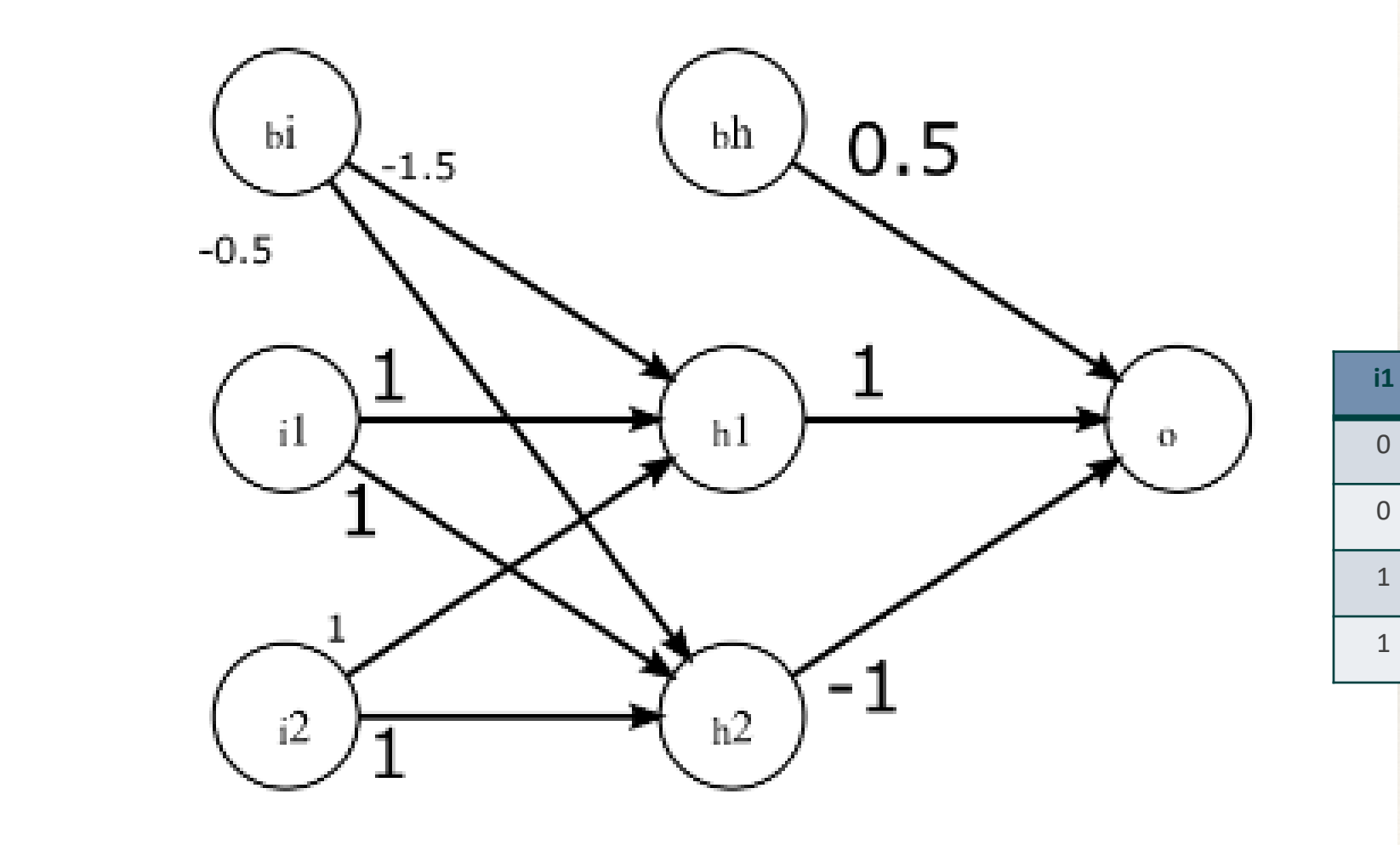
שאלה 8 - Feed forward (predict) –
רשת רב שכבתית לינארית – תרגיל (המשך)



איך תשנו את
הפרצפטרון כך
שיחשב את
פונקצית NXOR?

i1	i2
0	0
0	1
1	0
1	1

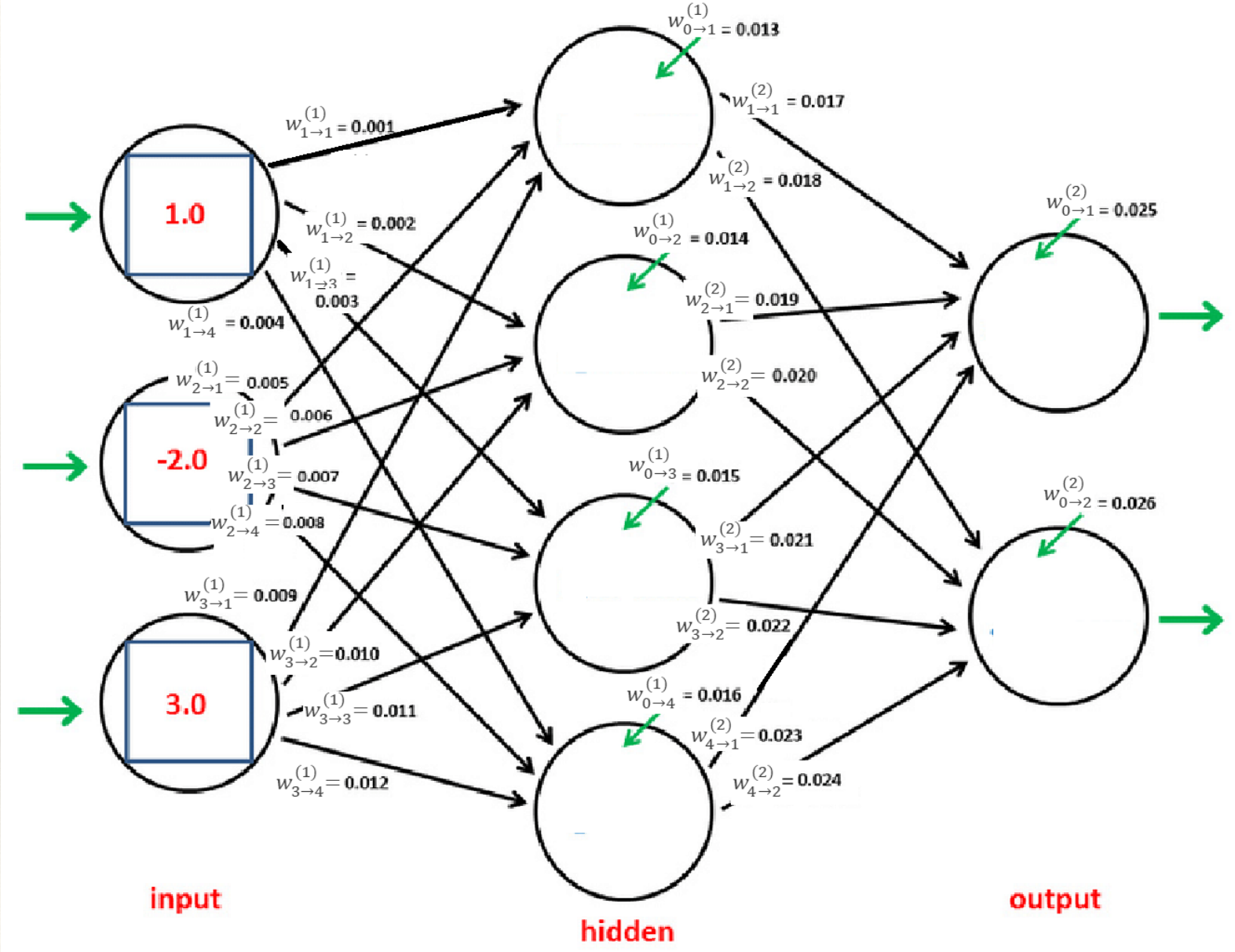
שאלה 8 - Feed forward (predict) –
רשת רב שכבתית לינארית – תרגיל (המשך) – פתרון



איך תשנו את הפרצפטרון כך שיחשב את פונקצית NXOR?

i1	i2	h1	h2	NXOR
0	0	-1.5 => -1	-0.5 => -1	.5 =>1
0	1	-0.5 => -1	0.5 => 1	-1.5 **=> -1**
1	0	-0.5 => -1	0.5 => 1	-1.5 **=> -1**
1	1	.5 => 1	1.5 = > 1	.5 => 1

שאלה 9א - Feed forward - (predict) - רשת רב שכבתית לינארית



מהי הטופולוגיה של רשת הנוירונים?

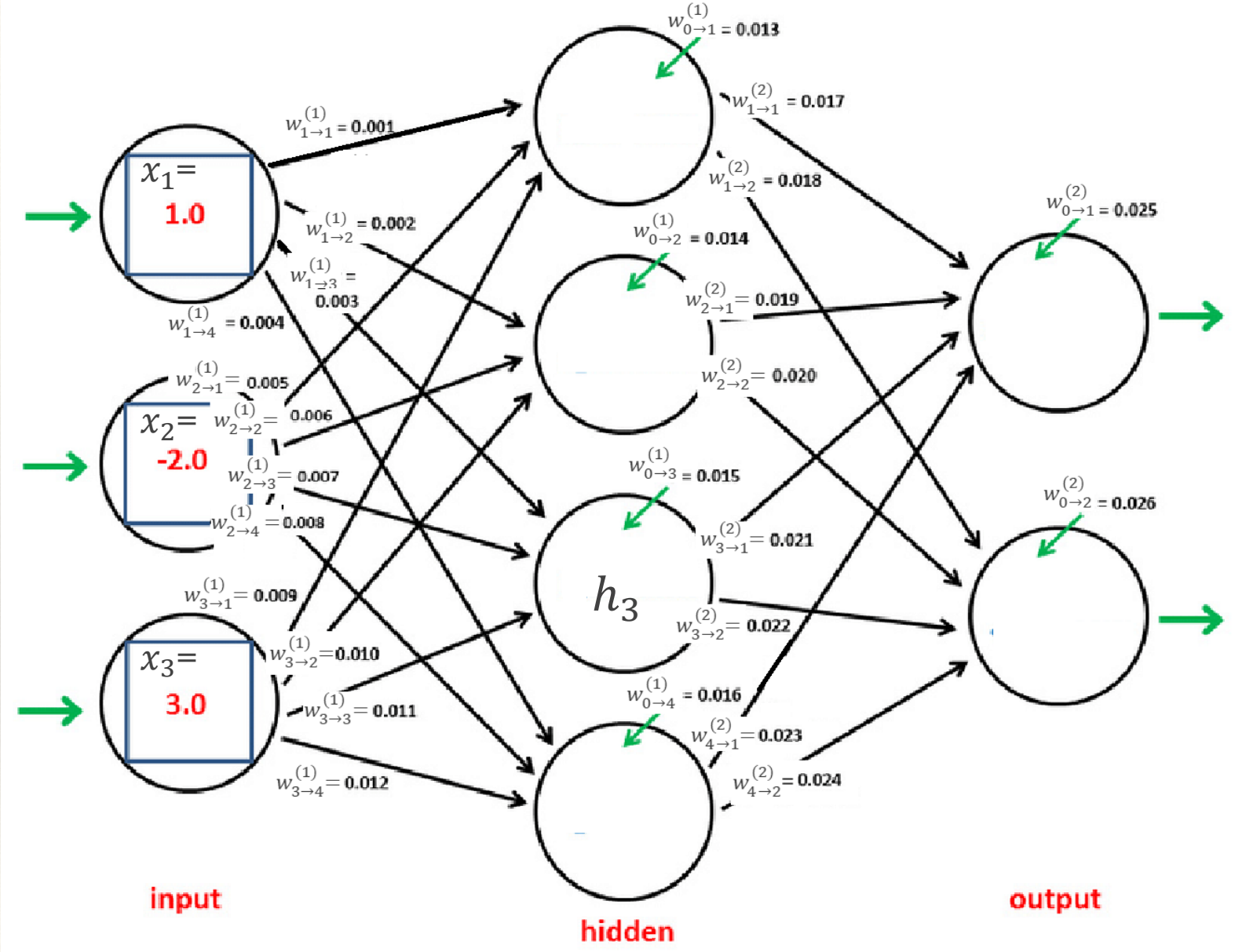
<u>תשובות אפשריות</u>:

1. 3 שכבות, 9 נוירונים (3 נוירונים התחלתיים, 4 חבויים, 2 בפלט)
2. 3 שכבות, 3 קלטים, 6 נוירונים (4 חבויים, 2 בפלט)
3. 4 שכבות – עליונה, אמצעיות ותחתונה (חבויה)

<u>תשובה</u>:

2. 3 שכבות, 3 קלטים, 6 נוירונים (4 חבויים, 2 בפלט)

שאלה 9ב – (predict) Feed forward - רשת רב שכבתית לינארית



מה יהיה ערכו של h_3 (הרשת לינארית, חשבו רק לפי הקומבינציה הלינארית, כלומר, פונקצית זהות כפונקצית אקטיבציה)?

<u>תשובות אפשריות:</u>

0.022 .1

2.0 .2 $(x_1+x_2+x_3=)$

0.037 .3

0.036 .4 $(w^{(1)}_{0\to3}+w^{(1)}_{1\to3}+w^{(1)}_{2\to3}+w^{(1)}_{3\to3}=)$

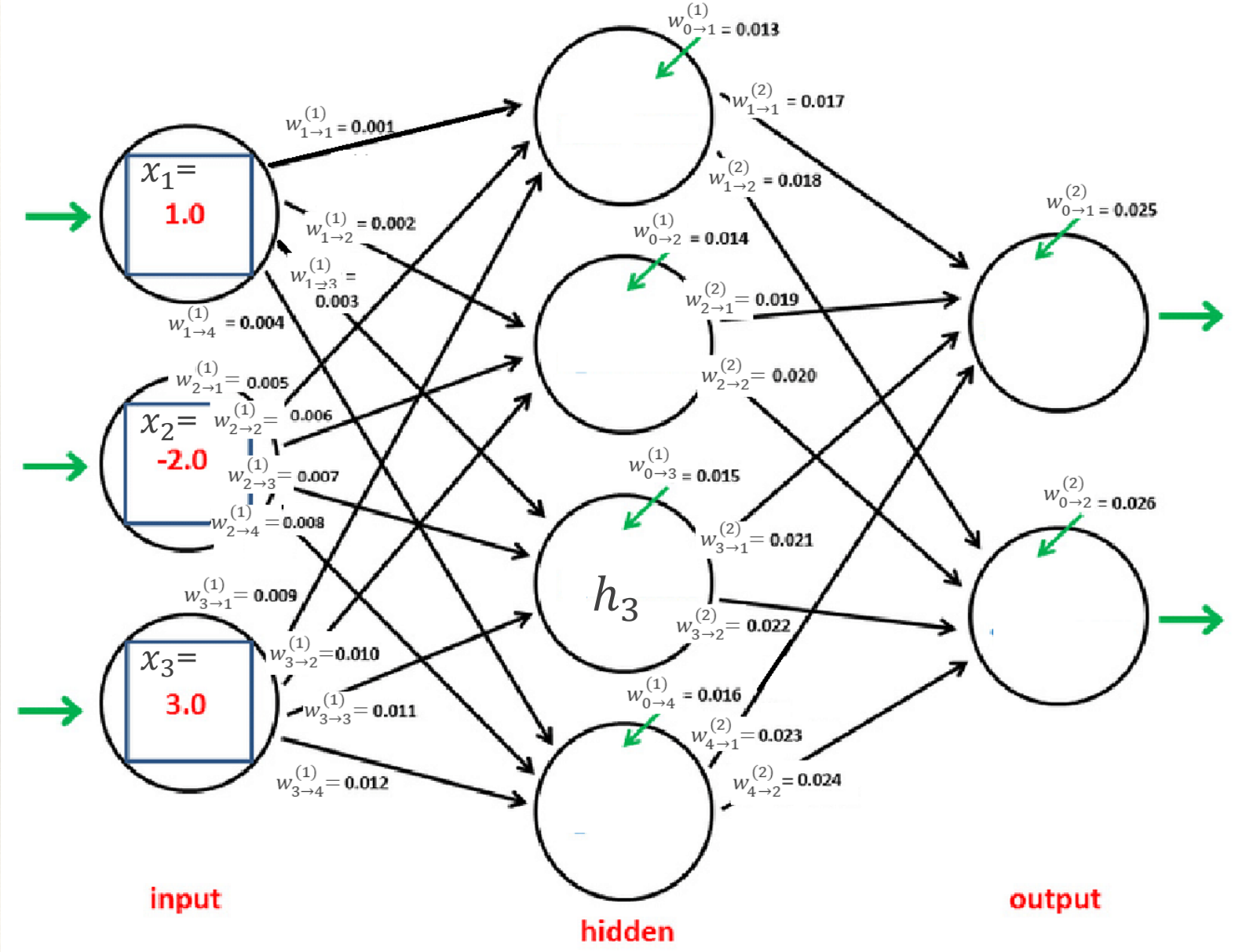
<u>תשובה:</u>

0.037 .3

<u>הסבר:</u>

0.022 $\left\{\begin{array}{l} 1.0*0.003\ (=x_1*w^{(1)}_{1\to3}) \\ -2.0*0.007\ (=x_2*w^{(1)}_{2\to3}) \\ +3.0*0.011\ (=x_3*w^{(1)}_{3\to3}) \\ +0.015\ (=w^{(1)}_{0\to3}) \end{array}\right.$ =0.037

שאלה 9ג - Feed forward (predict) - רשת רב שכבתית



מה יהיה ערכו של h_3 אם יחידות הנוירונים יהיו מסוג perceptron עם פונקצית sign?

<u>תשובות אפשריות:</u>

1. -1
2. 1
3. 0.037

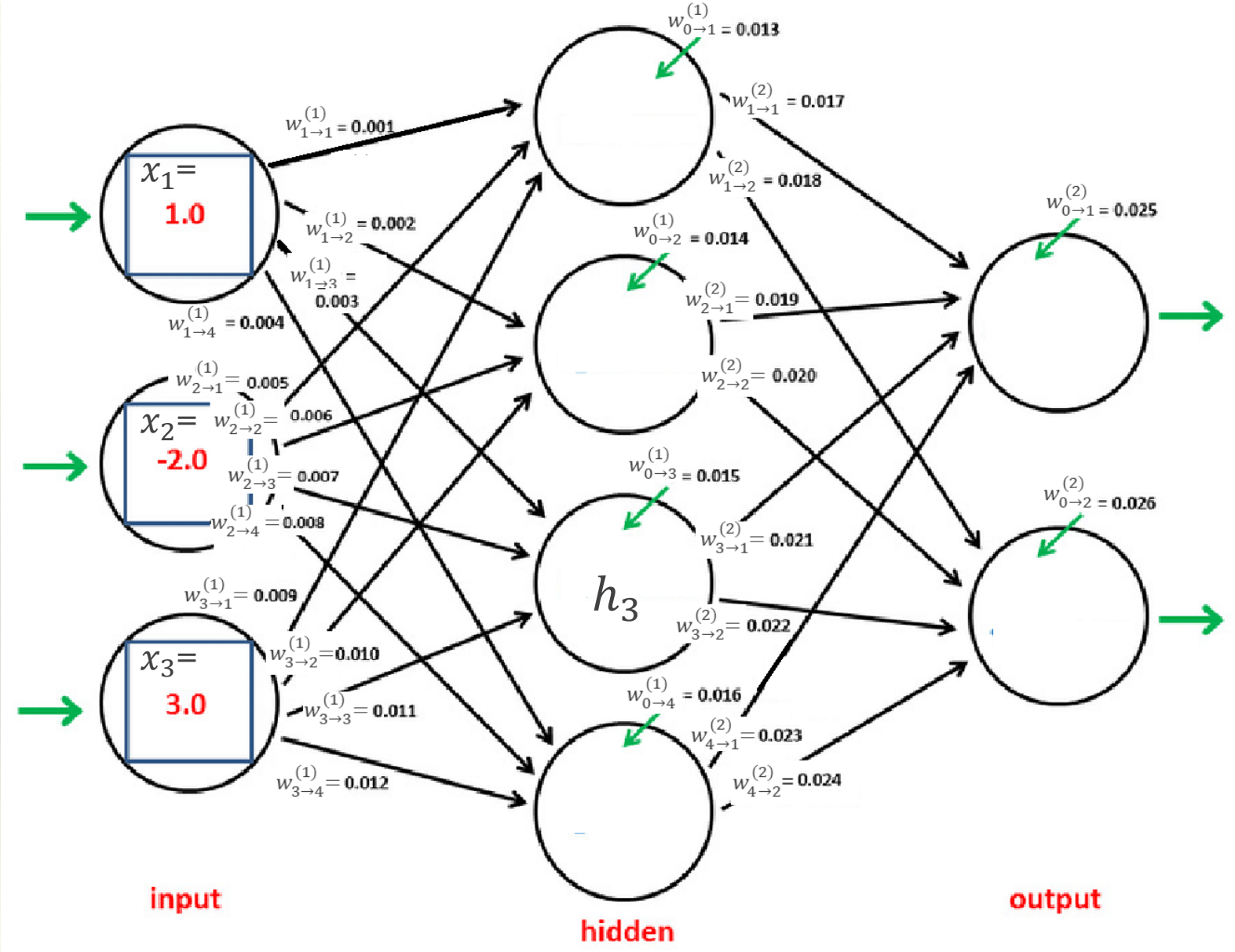
<u>תשובה:</u>

2. 1

<u>הסבר:</u> 0.037>0 ולכן התשובה 1

שאלה 9ד – (predict) Feed forward - רשת רב שכבתית

מה יהיה ערכו של h_3 אם יחידות הנוירונים יהיו מסוג נוירון עם פונקצית sigmoid?



sigmoid

$S(z) = \frac{1}{1+e^{-z}}$

תשובות אפשריות:

1. 0.509
2. 1
3. 0.037

תשובה:

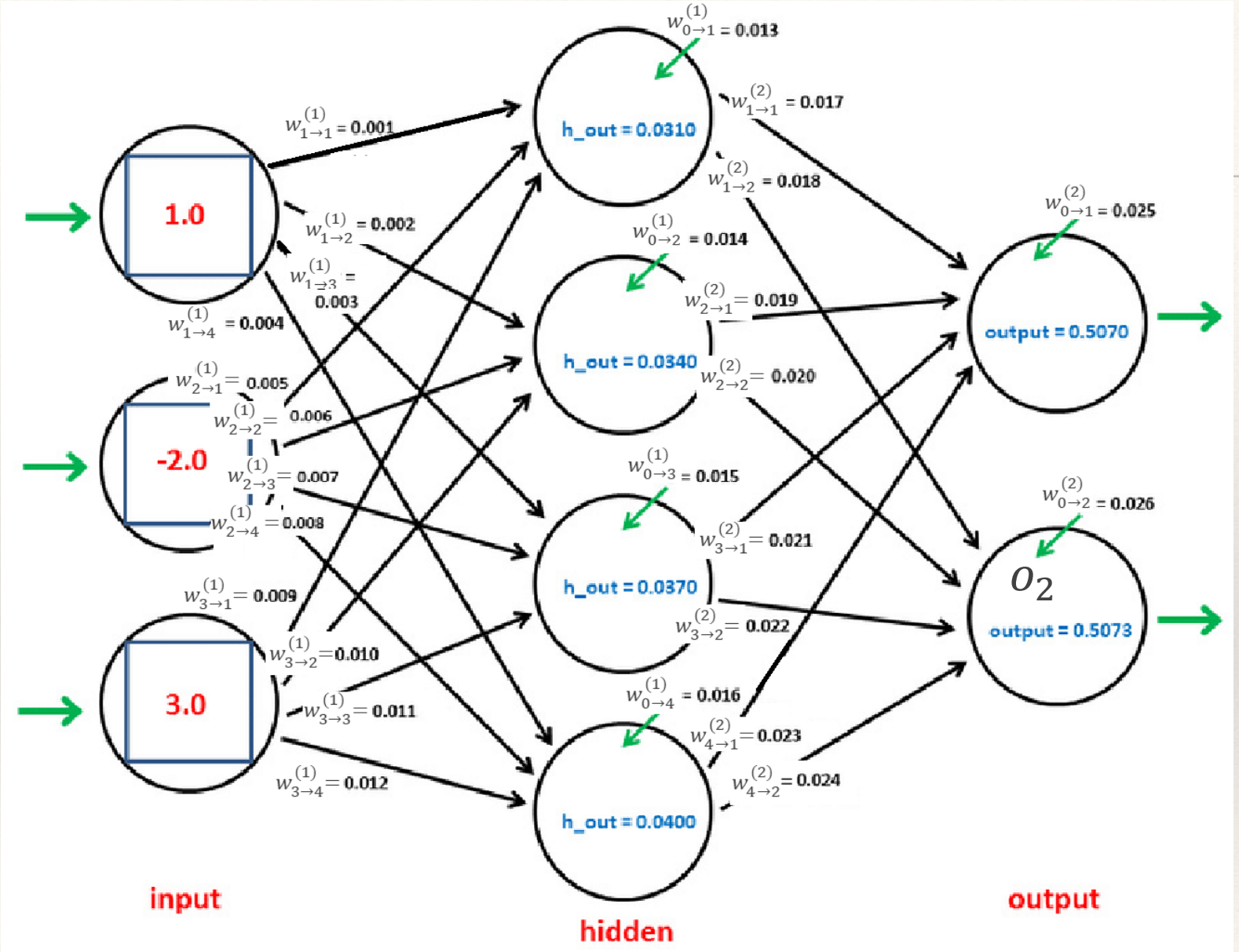
1. 0.509

הסבר:

$1/(1+e^{-net})=1/(1+e^{-0.037})$

$\approx 1/(1+0.9636) \approx 0.5092$

שאלה 9ה – (predict) Feed forward - רשת רב שכבתית לינארית



מה יהיה ערכו של o_2 (הרשת לינארית, חשבו רק לפי הקומבינציה הלינארית, כלומר, פונקצית זהות כפונקצית אקטיבציה)?

<u>תשובות אפשריות:</u>

1. 0.11
2. 0.029
3. 0.01
4. 0.126

<u>תשובה:</u>

2. 0.029

<u>הסבר:</u>

0.018*0.031+0.034*0.02+0.037*0.022+0.04*0.024+0.026≈0.029

Back Propagation

רשתות רב שכבתיות – אלגוריתם backpropagation - תזכורת

backpropagation(train_set, η, topology):

Initialization: Initialize all weights $(w_{i,j})$ to small random numbers.

Outer loop: for all train_set Until satisfied, Do

- For each training example $\langle \vec{x}, t_k \rangle$, Do

Feed forward:

1. Input the training example to the network and compute the network outputs (o_k)

Back propogation:

2. For each output unit k

$$\delta_k \leftarrow o_k(1-o_k)(t_k - o_k)$$

3. For each hidden unit h

$$\delta_h \leftarrow o_h(1-o_h) \sum_{k \in outputs} w_{h,k}\delta_k$$

4. Update each network weight $w_{i,j}$

$$w_{i,j} \leftarrow w_{i,j} + \Delta w_{i,j}$$

where

$$\Delta w_{i,j} = \eta \delta_j x_{i,j}$$

נתון כקלט לאלגוריתם:

Train set– כל זוגות האימון $\langle \vec{x}, t_k \rangle$.

$\vec{x}$ - מייצג את ה-feature vectors של האימון.

x_i - ערך מאפיין i (x_iים מהוויים את שכבת הקלט)

t_k - יכול להיות כמה ערכים (אחד עבור כל יחידה חיצונית).

η – קבוע הלמידה (מספר הקטן מ-1, גדול מ-0)

topology – המבנה של רשת הנוירונים

יחידה חיצונית:

o_k - פלט עבור יחידה חיצונית k

t_k - פלט מצופה עבור יחידה חיצונית k

δ_k - הטעות (בעזרת גרדיאנט) בשכבה חיצונית k

יחידה נסתרת:

o_h - פלט עבור יחידה נסתרת h

הערה: אין t_h - פלט מצופה עבור יחידה נסתרת h

δ_h - הטעות עבור יחידה נסתרת h (לפי התרומה היחסית ליחדות הפלט החיצוניות)

$w_{h,k}$ - המשקולת בין יחידה נסתרת h ליחידה חיצונית k

בכלל (i - קלט / יחידה נסתרת, j –יחידה נסתרת / חיצונית)

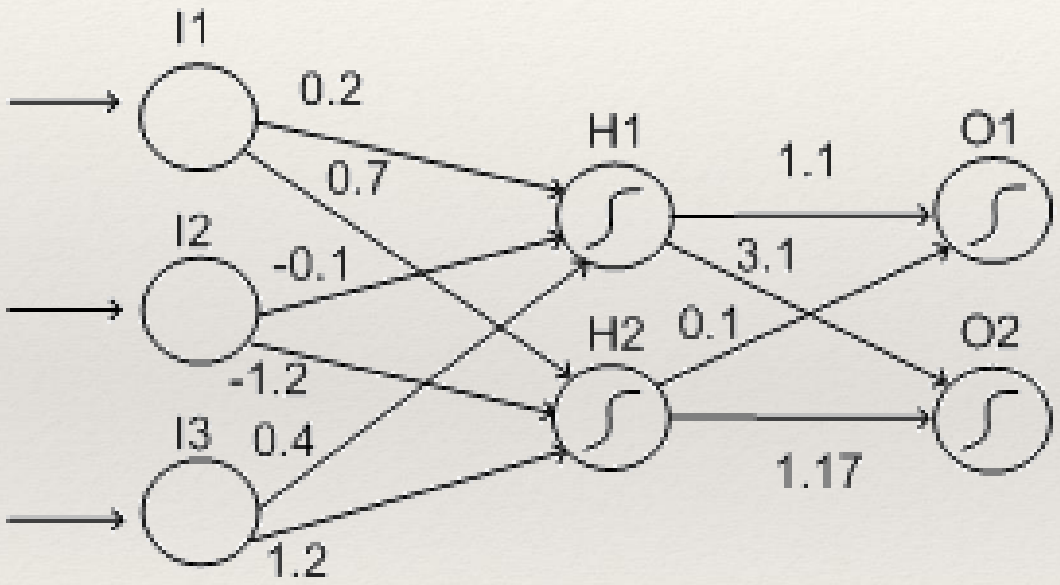
$w_{i,j}$ - משקולת הקשת בין i ל-j,

$\Delta w_{i,j}$ - התיקון למשקולת הקשת בין i ל- j

$x_{i,j}$ - הקלט המחובר לקשת $w_{i,j}$

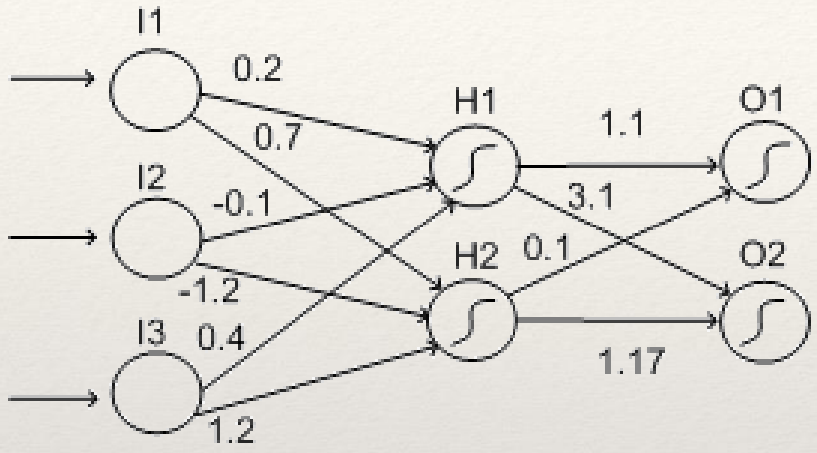
δ_j - הטעות בשכבה j

שאלה 10 – סימולציית backpropagation



ערוך סימולצית backpropagation
עבור הרשת הבאה, כאשר הדאטה סט
מורכב מדוגמא אחת
(10,30,20)
כאשר משתמשים ברשת לסיווג ל O_1 או
O_2 – לפי הערך הגבוה
במקרה שלנו הדוגמא שייכת ל O_1
(כלומר t1=1,t2=0)

שאלה 10 – סימולציית backpropagation –פתרון

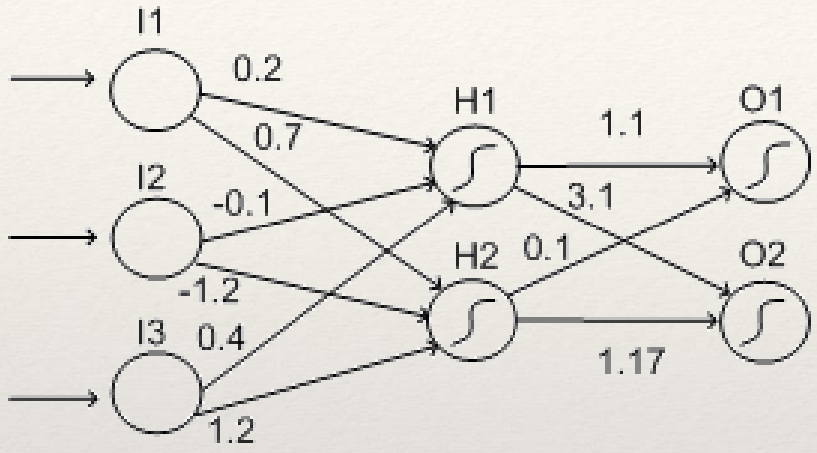


שלב I
Feedforward או Forward pass

דאטה סט:
(10,30,20)
output O1
(t1=1,t2=0 כלומר)

Input units		Hidden units			Output units		
Unit	Output	Unit	Weighted Sum Input	Output	Unit	Weighted Sum Input	Output
I1	10	H1	7	0.999	O1	1.0996	0.750
I2	30	H2	-5	0.0067	O2	3.1047	0.957
I3	20						

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



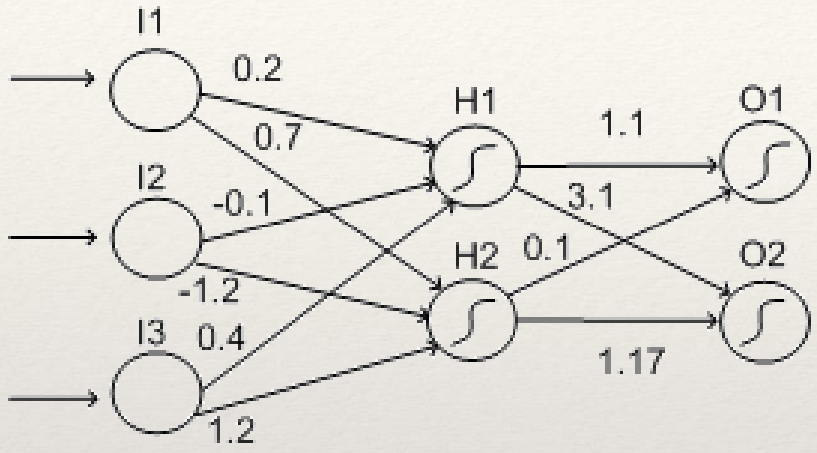
Input units		Hidden units			Output units		
Unit	Output	Unit	Weighted Sum Input	Output	Unit	Weighted Sum Input	Output
I1	10	H1	7	0.999	O1	1.0996	0.750
I2	30	H2	-5	0.0067	O2	3.1047	0.957
I3	20						

שלב I
Feedforward או Forward pass

דאטה סט:
(10,30,20)
output O1
(t1=1,t2=0 כלומר)

Output after this step is O2 (since O2>O1), which is wrong

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



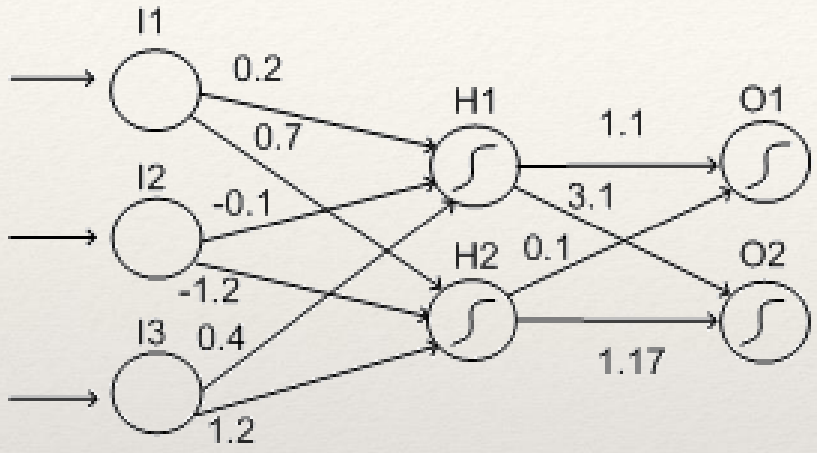
שלב II
Backpropagation

דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

מאחר והקטגוריה האמיתית היא o1 ז"א שהתוצאה של o1 היתה אמורה להיות 1, ושל o2 אמורה להיות 0, ולכן לפי הדוגמא

$t_1(E) = 1;$ $t_2(E) = 0;$ $o_1(E) = 0.750;$ $o_2(E) = 0.957$

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



שלב II
Backpropagation

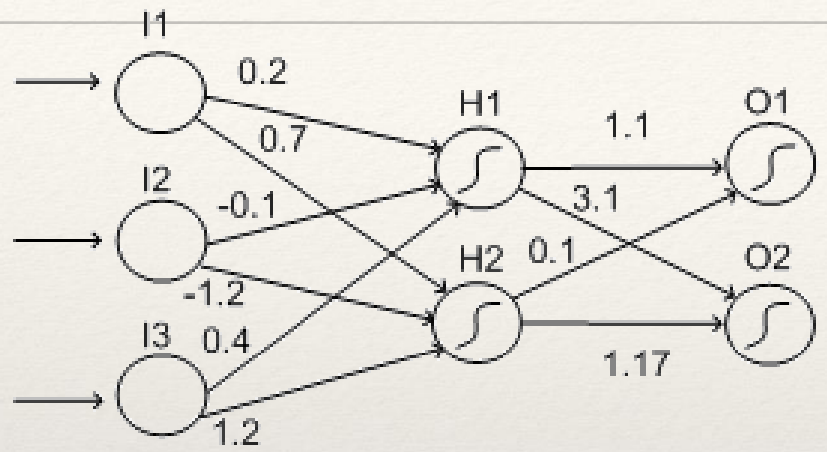
דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

נחשב את הטעות של יחידות הפלט:

$$\delta_{O1} = o_1(E)(1 - o_1(E))(t_1(E) - o_1(E)) = 0.750(1-0.750)(1-0.750) = 0.0469$$

$$\delta_{O2} = o_2(E)(1 - o_2(E))(t_2(E) - o_2(E)) = 0.957(1-0.957)(0-0.957) = -0.0394$$

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



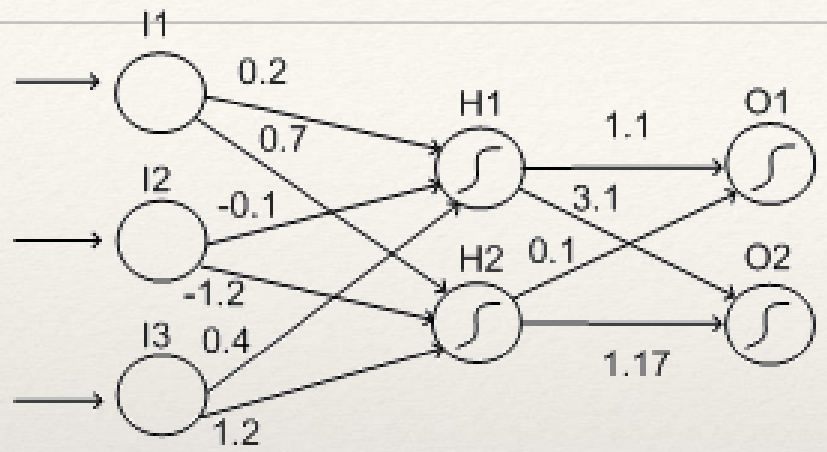
שלב II
Backpropagation

דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

ו"נפעפע" את המידע אחורה לשכבות הפנימיות. לצורך כך השגיאה של H1 נכפיל את את השגיאה של O1 במשקל מ H1 ל O1, ונוסיף זאת לשגיאה של של H1 כפול המשקל ל-O2. זה יוצא: (1.1*0.0469) + (3.1*-0.0394) = -0.0706.
כדי לחשב את השגיאה של H1 נכפיל את זה ב $h_1(E)*(1-h_1(E))$, כאשר $h_1(E)$ הינו הפלט של H1 עבור הדוגמא E (כמו בשלב ה FF) ונקבל:

$$\delta_{H1} = -0.0706*(0.999 * (1-0.999)) = -0.0000705$$

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



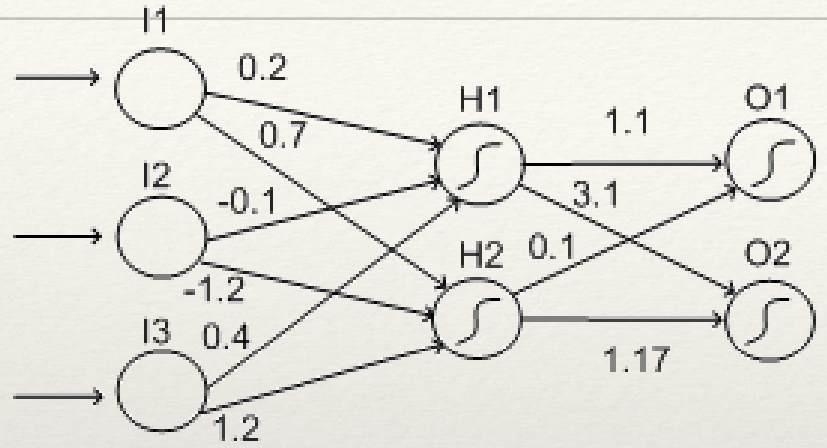
שלב II
Backpropagation

דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

ו"נפעפע" את המידע אחורה לשכבות הפנימיות. לצורך כך השגיאה של H2 נכפיל את את השגיאה של O1 במשקל מ H2 ל O1, ונוסיף זאת לשגיאה של של H2 כפול המשקל ל O2.
זה יוצא: (.1*0.0469) + (1.17*-0.0394) = -0.0414.
כדי לחשב את השגיאה של H2 נכפיל את זה ב $h_2(E)*(1-h_2(E))$, כאשר $h_2(E)$ הינו הפלט של H2 עבור הדוגמא E (כמו בשלב ה FF) ונקבל:

$$\delta_{H2}\ -0.0414 * (0.0067 * (1-0.0067)) = -0.000259$$

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



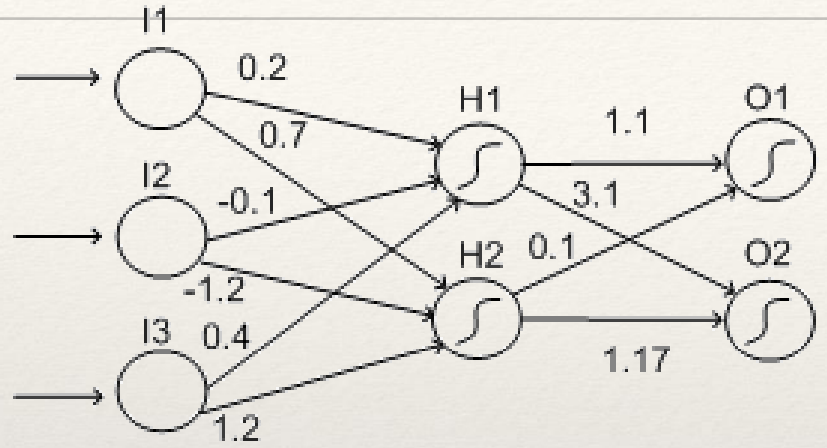
כעת יש את כל המידע כדי לחשב את שינוי המשקל עבור אטא=0.1 . נתחיל מ6 הערכים בין נוירוני הקלט לשכבת הביניים

Input unit	Hidden unit	η	δ_H	x_i	$\Delta = \eta * \delta_H * x_i$	Old weight	New weight
I1	H1	0.1	-0.0000705	10	-0.0000705	0.2	0.1999295
I1	H2	0.1	-0.00259	10	-0.00259	0.7	0.69741
I2	H1	0.1	-0.0000705	30	-0.0002115	-0.1	-0.1002115
I2	H2	0.1	-0.00259	30	-0.00777	-1.2	-1.20777
I3	H1	0.1	-0.0000705	20	-0.000141	0.4	0.39999
I3	H2	0.1	-0.00259	20	-0.00518	1.2	1.1948

שלב II
Backpropagation

דאטה סט:
(10,30,20)
output O1
(t1=1,t2=0 כלומר)

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



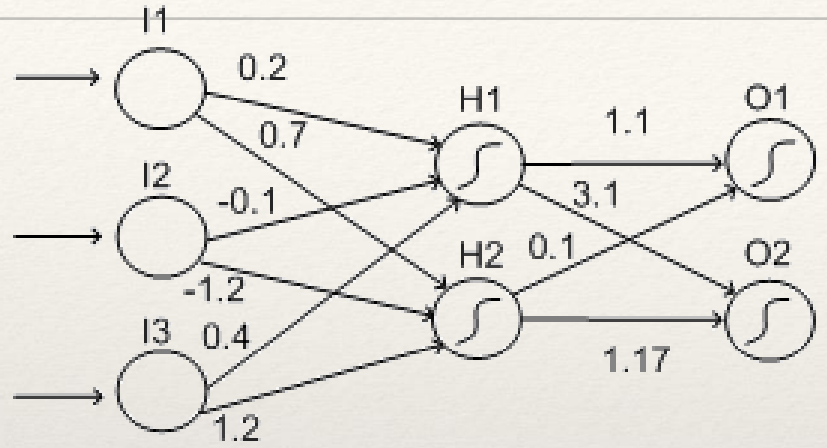
נעדכן את המשקולות לשכבת הפלט, באופן דומה, רק שבמקום להסתמך על ערכי הקלט של E נשתמש בערכי הסיגמויד שחושבו על ערכי שכבת הפלט (משלב ה-feed forward, בתחילת האיטרציה)

Hidden unit	Output unit	η	δ_O	$h_i(E)$	$\Delta = \eta * \delta_O * h_i(E)$	Old weight	New weight
H1	O1	0.1	0.0469	0.999	0.000469	1.1	1.100469
H1	O2	0.1	-0.0394	0.999	-0.00394	3.1	3.0961
H2	O1	0.1	0.0469	0.0067	0.00314	0.1	0.10314
H2	O2	0.1	-0.0394	0.0067	-0.0000264	1.17	1.16998

שלב II
Backpropagation

דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



שלב II
Backpropagation

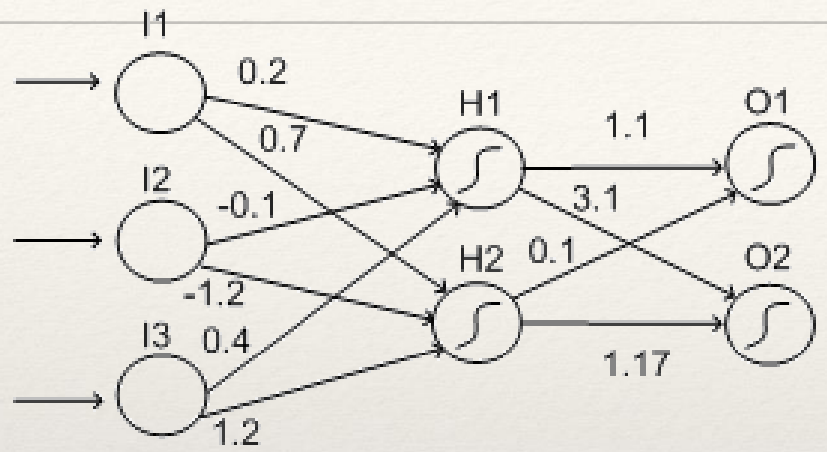
דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

ונמשיך לשכבת הביניים, באופן דומה, רק שבמקום להסתמך על ערכי הקלט של E נשתמש בערכי הסיגמויד שחושבו על ערכי שכבת הביניים

Hidden unit	Output unit	η	δ_O	$h_i(E)$	$\Delta = \eta * \delta_O * h_i(E)$	Old weight	New weight
H1	O1	0.1	0.0469	0.999	0.000469	1.1	1.100469
H1	O2	0.1	-0.0394	0.999	-0.00394	3.1	3.0961
H2	O1	0.1	0.0469	0.0067	0.00314	0.1	0.10314
H2	O2	0.1	-0.0394	0.0067	-0.0000264	1.17	1.16998

נשים לב שהערכים כמעט לא השתנו, כך שיתכן ובמקרה כזה כדאי להשתמש באטא (קבוע למידה) גדול יותר, אם כי כדאי לזכור שפונק סיגמויד, שינוי קטן בפונק בסכום המשקולות יכול ליצור שינוי גדול בפלט

שאלה 10 – סימולציית backpropagation – פתרון (המשך)



שלב II
Backpropagation

דאטה סט:
(10,30,20)
output O1
(כלומר t1=1,t2=0)

תרגיל בית: הריצו עוד איטרציה נוספת עם המשקולות החדשים..

Hidden unit	Output unit	η	δ_O	$h_i(E)$	$\Delta = \eta * \delta_O * h_i(E)$	Old weight	New weight
H1	O1	0.1	0.0469	0.999	0.000469	1.1	1.100469
H1	O2	0.1	-0.0394	0.999	-0.00394	3.1	3.0961
H2	O1	0.1	0.0469	0.0067	0.00314	0.1	0.10314
H2	O2	0.1	-0.0394	0.0067	-0.0000264	1.17	1.16998