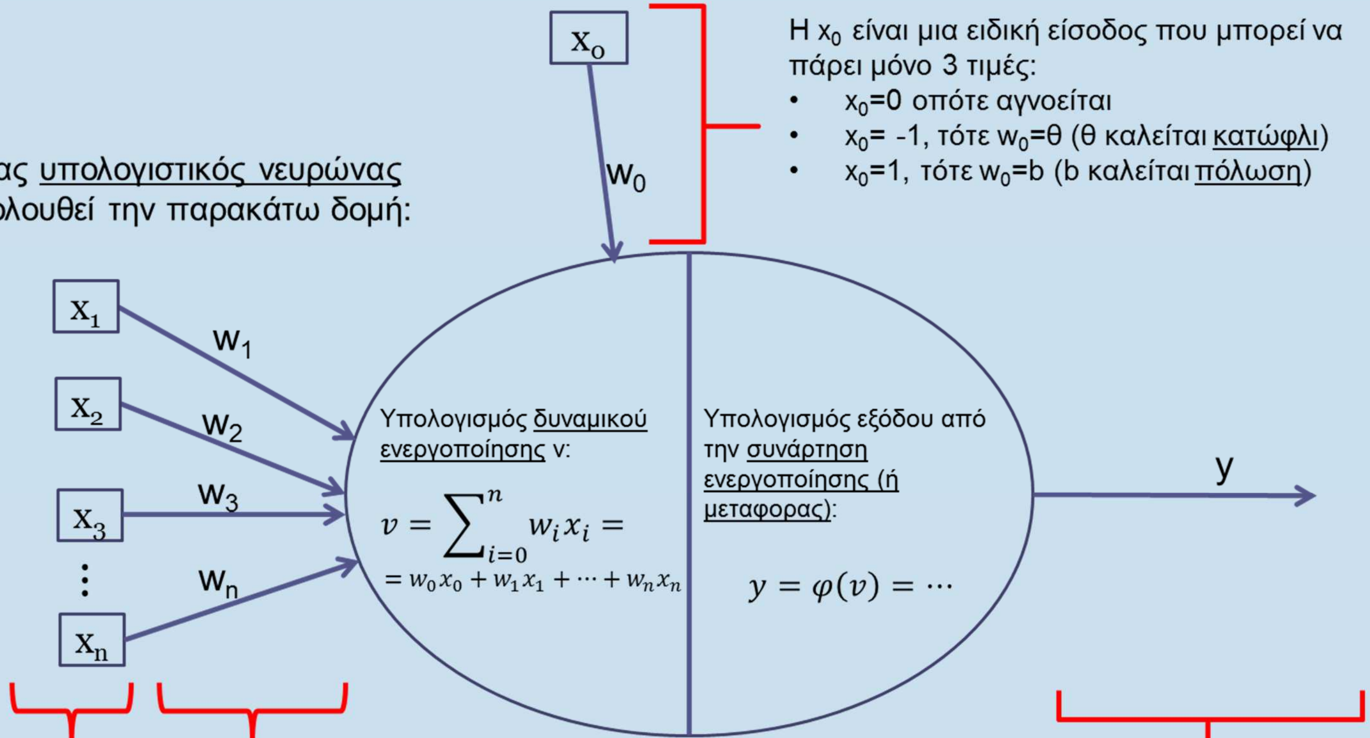
ΑΙΣΘΗΤΗΡΑΣ (ΤΡΟΠΟΣ ΛΕΙΤΟΥΡΓΙΑΣ)

ΛΕΙΤΟΥΡΓΙΑ ΕΝΟΣ ΝΕΥΡΩΝΑ

Ένας υπολογιστικός νευρώνας ακολουθεί την παρακάτω δομή:



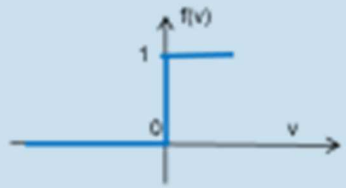
Η x_0 είναι μια ειδική είσοδος που μπορεί να πάρει μόνο 3 τιμές:

- $x_0=0$ οπότε αγνοείται
- $x_0= -1$, τότε $w_0=\theta$ (θ καλείται κατώφλι)
- $x_0=1$, τότε $w_0=b$ (b καλείται πόλωση)

ΣΥΝΑΡΤΗΣΕΙΣ ΕΝΕΡΓΟΠΟΙΗΣΗΣ

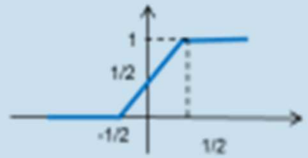
Βηματική Συνάρτηση

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases}$$



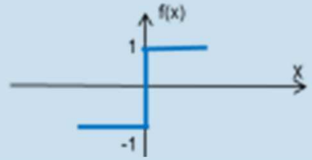
Τμηματικά Γραμμική Συνάρτηση

$$f(x) = \begin{cases} 1, & x \geq 0.5 \\ x, & -0.5 < x < 0.5 \\ 0, & x < -0.5 \end{cases}$$



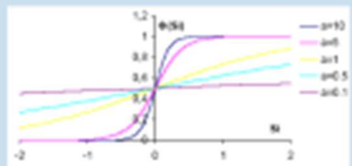
Συνάρτηση Προσήμου

$$f(x)=\text{sign}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$$



Σιγμοειδής Συνάρτηση

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



ΑΙΣΘΗΤΗΡΑΣ (ΕΚΠΑΙΔΕΥΣΗ ΜΕ ΓΡΑΦΙΚΗ ΕΠΙΛΥΣΗ)

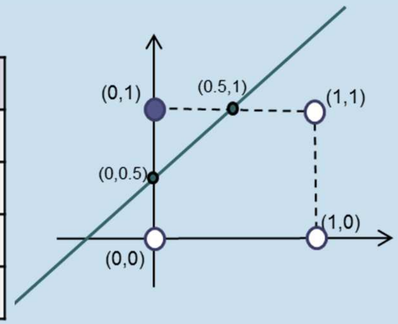
Παράδειγμα: Κατασκευάστε έναν αισθητήρα δύο εισόδων που ακολουθεί το μοντέλο McCullough-Pitts που αποφασίζει την λογική συνάρτηση:$X_2 \wedge \sim X_1$. Η επίλυση να γίνει με γραφική απεικόνιση της εξίσωσης ευθείας του νευρώνα.

<u>Βήμα 1:</u> Κατασκευάζουμε τον αληθοπίνακα σε σύστημα αξόνων (οριζόντιος άξονας το x1 και κάθετος άξονας το x2) τα σημεία κάνοντας μαύρα τα σημεία που είναι 1 και λευκά τα σημεία που είναι 0.

<u>Βήμα 2:</u> Σχεδιάζουμε μια ευθεία που διαχωρίζει τα πρότυπα των δύο κλάσεων, έτσι ώστε να περνάει από δύο συγκεκριμένα σημεία των όποιων οι συντεταγμένες είναι εύκολο να εντοπιστούν. Ειδικά για λογικές πύλες, οι συντεταγμένες των σημείων θα είναι πολλαπλάσια του 0.5

Επίλυση: έχουμε:

x_1	x_2	Έξοδος: $X_2 \wedge \sim X_1$
0	0	0
0	1	1
1	0	0
1	1	0



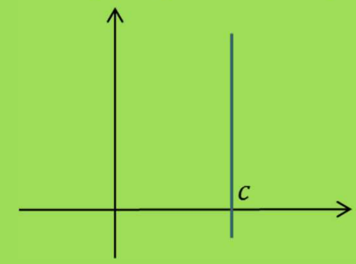
<u>Βήμα 3:</u> Βρίσκουμε την ευθεία απόφασης ως εξής. Ονομάζουμε τα δύο σημεία (x1,y1) και (x2,y2) και υπολογίζουμε την εξίσωση ευθείας από τον τύπο: $\frac{x-x_1}{x_1-x_2}=\frac{y-y_1}{y_1-y_2}$.

Έπειτα φέρουμε την εξίσωση ευθείας στη μορφή: $\alpha x+\beta y+\gamma=0$

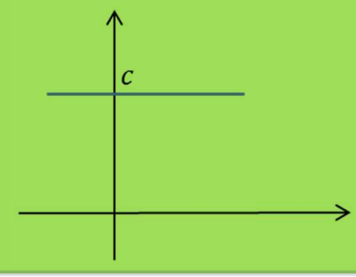
Επίλυση: Δύο σημεία από τα οποία διέρχεται η ευθεία είναι:
$(x_1,y_1)=(0{,}0.5)$ και $(x_2,y_2)=(0.5{,}1)$
Άρα η ζητούμενη ευθεία είναι η:

$$\frac{x-x_1}{x_1-x_2}=\frac{y-y_1}{y_1-y_2}\Rightarrow\frac{x-0}{0-0.5}=\frac{y-0.5}{0.5-1}\Rightarrow\frac{x}{-0.5}=\frac{y-0.5}{-0.5}\Rightarrow$$
$$-0.5x=-0.5(y-0.5)\Rightarrow -0.5x=-0.5y+0.25\Rightarrow(-0.5)x+0.5y-0.25=0$$

(Ειδικές Περιπτώσεις Ευθειών)



Ευθεία: $x=c$
και γράφεται:
$1x+0y+(-c)=0$



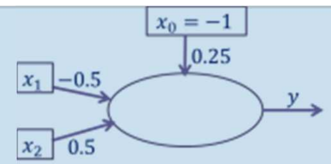
Ευθεία: $y=c$
και γράφεται:
$0x+1y+(-c)=0$

<u>Βήμα 4:</u> Κάνουμε 1:1 συσχέτιση των σταθερών των εξισώσεων:

Εξίσωση Ευθείας: $\alpha x\ +\ \beta y\ +\gamma=0$

Εξίσωση Νευρώνα: $w_1x_1+w_2x_2-\theta=0$

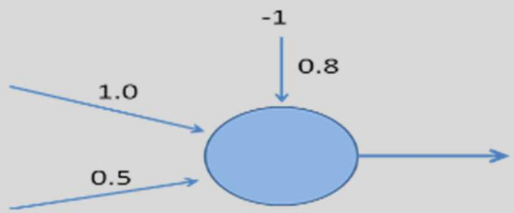
Επίλυση:
Εξίσωση Ευθείας: $(-0.5)x+0.5y-0.25=0$
Εξίσωση Νευρώνα: $w_1x_1+w_2x_2-\theta=0$
Συνεπώς τα βάρη του νευρώνα είναι: $w_1=-0.5, w_2=0.5, \theta=0.25$



Αρχικοποίηση:

- Αρχικοποιούμε τα διανύσματα:
  - $W = [w_0, w_1, \dots, w_n]$ με τα αρχικοποιημένα βάρη των ακμών
  - Για κάθε πρότυπο 1,…,K: Κατασκευάζουμε το διάνυσμα: $x_i = [x_{i0}, x_{i1}, \dots, x_{in}]$ και αρχικοποιούμε την επιθυμητή έξοδο: d_i
- Δίνουμε τιμή στην παράμετρο μάθησης η: 0<η<1

Εκφώνηση: Θέλουμε να εκπαιδεύσουμε έναν αισθητήρα ώστε να επιλύει το πρόβλημα του λογικού ΚΑΙ. Κατά τη διάρκεια της διαδικασίας εκπαίδευσης με ρυθμό εκπαίδευσης n=0.5, έχουμε τα βάρη που φαίνονται στο Σχήμα. Ο αισθητήρας ακολουθεί το μοντέλο McCullough-Pits με εξόδους 1 και 0.



Να συνεχίσετε τη διαδικασία εκπαίδευσης έως ότου να εκπαιδευτεί ο αισθητήρας, παρουσιάζοντας διαδοχικά τα διανύσματα I_1=(0,0), I_2=(0,1), I_3=(1,0) και I_4=(1,1).

Αρχικοποίηση:

Κωδικοποιήση των εισόδων ως διανύσματα με την επιθυμητή έξοδο:

- Είσοδος: $x_1 = [-1,0,0]^T$ Επιθυμητή Έξοδος: $d_1 = 0$
- Είσοδος: $x_2 = [-1,0,1]^T$ Επιθυμητή Έξοδος: $d_2 = 0$
- Είσοδος: $x_3 = [-1,1,0]^T$ Επιθυμητή Έξοδος: $d_3 = 0$
- Είσοδος: $x_4 = [-1,1,1]^T$ Επιθυμητή Έξοδος: $d_4 = 1$

Αρχικοποίηση των αρχικών βαρών ως διάνυσμα: $W = [0.8,1.0,0.5]^T$

Ρυθμός Εκπαίδευσης: $\eta = 0.5$

Κύκλος Εκπαίδευσης: Για κάθε πρότυπο: $i = 1 \ldots . K$:

- Υπολόγισε το δυναμικό για το πρότυπο i ως: $\mathrm{v} = W^{\mathrm{T}} \cdot x_i$
- Υπολόγισε την έξοδο από την συνάρτηση δυναμικού: $y_i = \varphi(\mathrm{v})$
- Υπολόγισε το σφάλμα ως: $e = d_i - y_i$
- Αν το σφάλμα δεν είναι μηδενικό
  - Υπολογίζονται νέα βάρη ως: $W_{new} = W_{old} + \eta * e * x_i$ (**Κανόνας Δέλτα**)

1ο πρότυπο $x_1 = [-1{,}0{,}0],\ d_1 = 0$

- $W^{\mathrm{T}} \cdot x_1 = [0.8{,}1.0{,}0.5] \times \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -0.8$
- Αφού $-0.8 < 0$ η έξοδος είναι $y_1 = 0$.
- Σφάλμα: $\text{error} = d_1 - y_1 = 0$. Τα βάρη δεν αλλάζουν.

2ο πρότυπο $x_2 = [-1{,}0{,}1],\ d_2 = 0$

- $W^{\mathrm{T}} \cdot x_2 = [0.8{,}1.0{,}0.5] \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = -0.3$
- Αφού $-0.3 < 0$ η έξοδος είναι $y_2 = 0$.
- Σφάλμα: $\text{error} = d_2 - y_2 = 0$. Τα βάρη δεν αλλάζουν.

3ο πρότυπο $x_3 = [-1{,}1{,}0],\ d_3 = 0$

- $W^{\mathrm{T}} \cdot x_3 = [0.8{,}1.0{,}0.5] \times \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = -0.8 + 1.0 = 0.2$
- Αφού $0.2 \geq 0$ η έξοδος είναι $y_3 = 1$.
- Σφάλμα: $\text{error} = d_3 - y_3 = -1$. Τα βάρη αλλάζουν
- $W = W + \eta \cdot error \cdot x_3 =$

$$= \begin{bmatrix} 0.8 \\ 1.0 \\ 0.5 \end{bmatrix} + 0.5 * (-1) * \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.0 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0.5 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.3 \\ 0.5 \\ 0.5 \end{bmatrix}$$

4ο πρότυπο $x_4 = [-1{,}1{,}1],\ d_4 = 1$

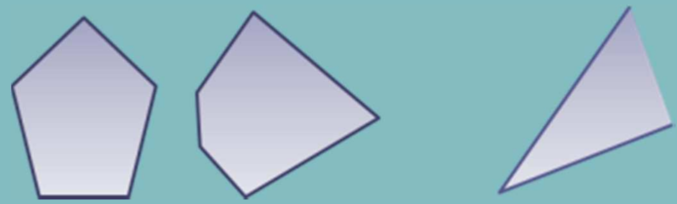
- $W^{\mathrm{T}} \cdot x_4 = [1.3{,}0.5{,}0.5] \times \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = -1.3 + 0.5 + 0.5 = -0.3$
- Αφού $-0.3 < 0$ η έξοδος είναι $y_4 = 0$.
- Σφάλμα: $\text{error} = d_4 - y_4 = 1$. Τα βάρη αλλάζουν
- $W = W + \eta \cdot error \cdot x_3 =$

$$= \begin{bmatrix} 1.3 \\ 0.5 \\ 0.5 \end{bmatrix} + 0.5 * 1 * \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.3 \\ 0.5 \\ 0.5 \end{bmatrix} + \begin{bmatrix} -0.5 \\ 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.0 \\ 1.0 \end{bmatrix}$$

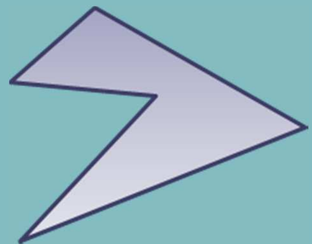
Ολοκλήρωση Κύκλου Εκπαίδευσης:

- Ο νευρώνας δεν απάντησε σωστά σε όλα τα πρότυπα, άρα θα πραγματοποιηθεί και άλλος κύκλος εκπαίδευσης.

Κυρτές Περιοχές:

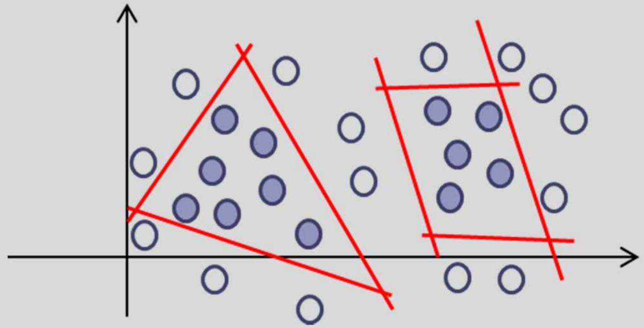


Μη Κυρτές Περιοχές:



Η μεθοδολογία κατασκευάζει ΤΝΔ που αποφασίζει την ένωση κυρτών περιοχών

Εκφώνηση: Δώστε τοπολογία ΤΝΔ που αποφασίζει το ακόλουθο πρόβλημα:



- Δίνονται μόνο τα πρότυπα (σημεία) στα οποία προσδιορίζεται η κλάση τους (λευκό η μαύρο χρώμα)
- «Εμφωλιάζουμε» τα μαύρα πρότυπα σε περιοχές απόφασης (μέρη του επιπέδου που ορίζονται από ευθείες έτσι ώστε τα πρότυπα να ανήκουν σε μόνο μία κλάση)

Λύση:

Επίπεδο Εισόδου:
- 2 (μη υπολογιστικοί) νευρώνες (αφού είμαστε στο επίπεδο)

Γενίκευση: Όσες και οι διαστάσεις των δεδομένων εισόδου

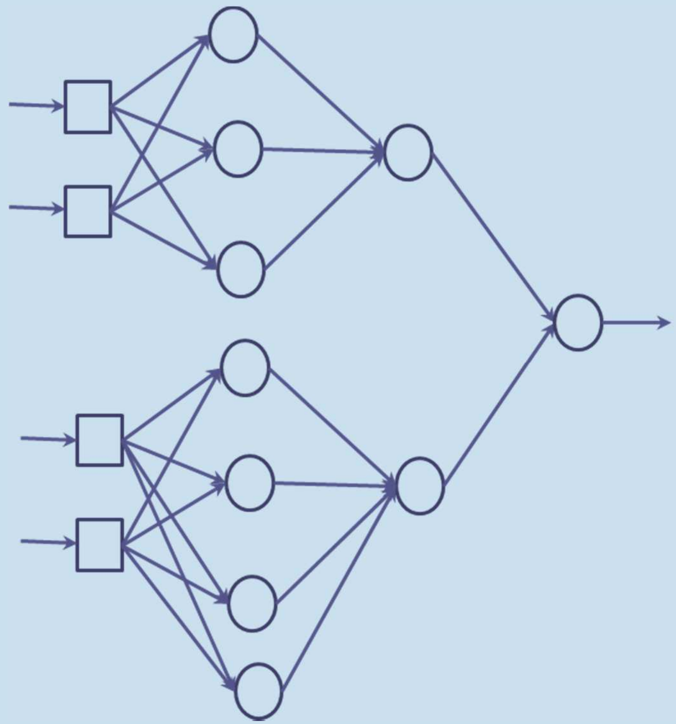
Κρυφό Επίπεδο:
- Ένας νευρώνας για κάθε ευθεία απόφασης

Αντίστοιχα για 3διάστατα δεδομένα => υπερεπίπεδο απόφασης

Κρυφό Επίπεδο 2:
- Ένας νευρώνας για κάθε περιοχή απόφασης (υλοποιεί το λογικό AND των αντίστοιχων ευθειών που ορίζουν την περιοχή απόφασης)

Επίπεδο Εξόδου:
- Ένας Νευρώνας (υλοποιεί το λογικό OR των περιοχών απόφασης)

2 κλάσεις => 1 νευρώνας εξόδου
3 ή 4 κλάσεις => 2 νευρώνες εξόδου
5 ή 6 ή 7 ή 8 κλάσεις => 3 νευρώνες εξόδου …κ.ο.κ.

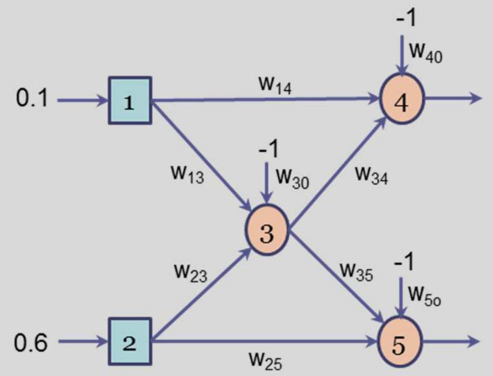


Αρχικοποίηση:

- Αρχικοποιούμε τα διανύσματα:
  - Για κάθε πρότυπο 1,…,K: Κατασκευάζουμε το διάνυσμα: $x_i = [x_{i0}, x_{i1}, \dots, x_{in}]$ και αρχικοποιούμε την επιθυμητή έξοδο: d_i
- Δίνουμε αρίθμηση στους κόμβους (αν αυτή δεν υπάρχει ήδη)
  - Πρέπει να υπάρχει μία τοπολογική ταξινόμηση στους κόμβους (δηλαδή να μην υπάρχει ακμή από κόμβο σε προηγούμενό του κόμβο)
- Αρχικοποιούμε τις τιμές των βαρών σύμφωνα με την εκφώνηση.
- Εντοπίζουμε την συνάρτηση ενεργοποίησης για κάθε κόμβο καθώς και την παραγωγό της (θα είναι κάποια συνεχής συνάρτηση)
- Δίνουμε τιμή στην παράμετρο μάθησης η: 0<η<1 (από εκφώνηση)

Πραγματοποιούμε κύκλους εκπαίδευσης διαδοχικά για τα πρότυπα.

Εκφώνηση: Δίνεται ένα πολυεπίπεδο ΤΝΔ τοπολογίας 2-1-2 με τη συνδεσμολογία όπως φαίνεται στο παρακάτω σχήμα.
Για την εκπαίδευσή του χρησιμοποιείται η μέθοδος οπισθοδιάδοσης του σφάλματος με ρυθμό εκπαίδευσης n=1, χωρίς χρήση ορμής (momentum).
Η συνάρτηση ενεργοποίησης σε όλους τους νευρώνες είναι η σιγμοειδής συνάρτηση S, όπου: $S(x) = \frac{1}{1+e^{-x}}$



Πίνακας 1			
Βάρος	Τιμή	Βάρος	Τιμή
w_{13} =	0,5	$w_{30} = \theta_3$	0,4
w_{14} =	0,5	$w_{40} = \theta_4$	0,4
w_{23} =	0,4	$w_{50} = \theta_5$	0,4
w_{25} =	0,4		
w_{34} =	0,3		
w_{35} =	0,3		

Σε κάποια στιγμή εκπαίδευσής του για την εκμάθηση του προτύπου [0.1,0.6] με επιθυμητή έξοδο [0.0, 1.0] τα βάρη των συνδέσεων και οι τιμές των κατωφλίων έχουν πάρει τις τιμές που δίνονται στον Πίνακα 1. Θεωρείστε ότι τα κατώφλια είναι συνάψεις με είσοδο –1 και βάρος ίσο με την τιμή του κατωφλίου. Να κάνετε τις πράξεις με ακρίβεια 3 δεκαδικών ψηφίων.

Να πραγματοποιήσετε έναν πλήρη κύκλο εκπαίδευσης (προς τα εμπρός και προς τα πίσω πέρασμα)

Συνεπώς:
Χρησιμοποιείται ένα πρότυπο εισόδου

Είσοδος:	Επιθυμητή Έξοδος:
x_1=0.1	d_4=0.0
x_2=0.6	d_5=1.0

ΠΡΟΣ ΤΑ ΕΜΠΡΟΣ ΠΕΡΑΣΜΑ:

Οι νευρώνες εξετάζονται κατά την αύξουσα αρίθμηση: j=1….N

- Για κάθε νευρώνα εισόδου θέτουμε ως y_j την είσοδο που παράγει.
- Για κάθε υπολογιστικό νευρώνα j (κρυφό και εξόδου):

  Υπολόγισε το δυναμικό ως: $v_j = \sum_{i=0}^{p} w_{ij}\, y_i$

  Υπολόγισε την έξοδο από την συνάρτηση ενεργοποίησης: $y_j = \varphi(v_j)$
- Συμβολίζουμε με o_j την έξοδο μόνο των νευρώνων εξόδου

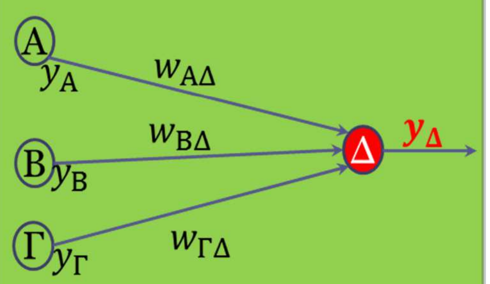
  $o_j = y_j$
- Για κάθε νευρώνα εξόδου:

  Υπολόγισε το σφάλμα: $e_j = d_j - o_j$ (επιθυμητή μείον παραγματική)

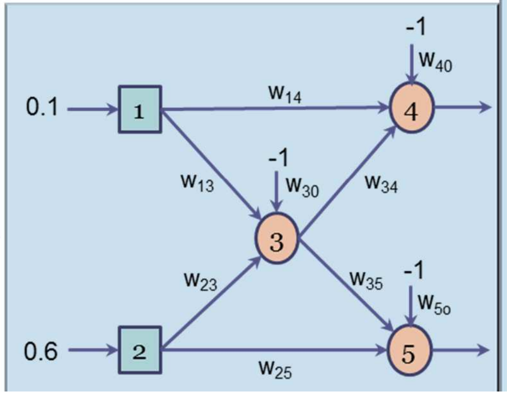
p είναι ο συνολικός αριθμός εισόδων του νευρώνα j

Υπολογίζεται το δυναμικό του νευρώνα ως άθροισμα των γινομένων βαρών-εισόδων

Συμπεριλαμβάνεται η είσοδος κατωφλίου (αν υπάρχει)



$v_\Delta = w_{A\Delta} y_A + w_{B\Delta} y_B + w_{\Gamma\Delta} y_\Gamma$
$y_\Delta = \varphi(v_\Delta)$



ΠΡΟΣ ΤΑ ΕΜΠΡΟΣ ΠΕΡΑΣΜΑ

ΝΕΥΡΩΝΑΣ 1 (νευρώνας εισόδου)
Η είσοδος μεταφέρεται στην έξοδο, άρα $y_1=x_1=0.1$

ΝΕΥΡΩΝΑΣ 2 (νευρώνας εισόδου)
Η είσοδος μεταφέρεται στην έξοδο, άρα $y_2=x_2=0.6$

ΝΕΥΡΩΝΑΣ 3 (Κρυφός Νευρώνας)
Δυναμικό: $v_3 = (w_{13} \cdot y_1) + (w_{23} \cdot y_2) + (w_{30} \cdot (-1)) = (0{,}5 \cdot 0{,}1) + (0{,}4 \cdot 0{,}6) + (0{,}4 \cdot (-1)) = -0.11$
Ενεργοποίηση: $y_3 = \varphi(v_3) = \frac{1}{1+e^{-(-0.11)}} = 0.473$

ΝΕΥΡΩΝΑΣ 4 (Νευρώνας Εξόδου)
Δυναμικό: $v_4 = (w_{14} \cdot y_1) + (w_{34} \cdot y_3) + (w_{40} \cdot (-1)) = (0{,}5 \cdot 0{,}1) + (0{,}3 \cdot 0{,}473) + (0{,}4 \cdot (-1)) = -0.208$
Ενεργοποίηση: $y_4 = \varphi(v_4) = \frac{1}{1+e^{-(-0.208)}} = 0.448$

ΝΕΥΡΩΝΑΣ 5 (Νευρώνας Εξόδου)
Δυναμικό: $v_5 = (w_{35} \cdot y_3) + (w_{25} \cdot y_2) + (w_{50} \cdot (-1)) = (0{,}3 \cdot 0.473) + (0{,}4 \cdot 0{,}6) + (0{,}4 \cdot (-1)) = -0.018$
Ενεργοποίηση: $y_5 = \varphi(v_5) = \frac{1}{1+e^{-(-0.018)}} = 0.496$

Συνεπώς η έξοδος των νευρώνων είναι:

y_1	y_2	y_3	y_4	y_5
0.1	0.6	0.473	0.448	0.496

Υπολογισμός Σφάλματος για τους νευρώνες εξόδου:
Νευρώνας 4: $e_4 = d_4 - y_4 = 0 - 0.448 = -0.448$
Νευρώνας 5: $e_5 = d_5 - y_5 = 1 - 0.496 = 0.504$

Άρα τα σφάλματα στους νευρώνες εξόδου είναι:

e_4	e_5
−0.448	0.504

ΠΡΟΣ ΤΑ ΠΙΣΩ ΠΕΡΑΣΜΑ:
Α. ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΠΙΚΩΝ ΚΛΙΣΕΩΝ

Οι νευρώνες εξετάζονται κατά την φθίνουσα αρίθμηση j=N, N-1,…,1

Υπολογισμός της τοπικής κλίσης δ για κάθε υπολογιστικό νευρώνα:

Για τους νευρώνες εξόδου

$$\delta_j(n) = e_j \cdot \varphi'_j(v_j)$$

Για τους νευρώνες κρυφού επιπέδου:

$$\delta_j(n) = \varphi'_j(v_j) \cdot \sum_k [\delta_k(n) \cdot w_{jk}(n)]$$

Για τους νευρώνες εισόδου: Δεν γίνεται υπολογισμός τοπικής κλίσης

Β. ΔΙΟΡΘΩΣΕΙΣ ΣΤΑ ΒΑΡΗ ΤΩΝ ΑΚΜΩΝ

Διορθώσεις σε όλα τα βάρη:

Υπολογισμός Διόρθωσης των Βαρών των ακμών:

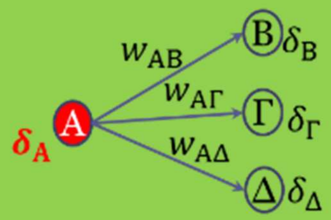
$$\Delta w_{ij}(n) = \eta \cdot \delta_j(n) \cdot y_i(n)$$

Υπολογισμός των βαρών:

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n)$$

Το δ για τους νευρώνες εξόδου υπολογίζεται ως το γινόμενο
(Σφάλμα του νευρώνα)x
(Την παράγωγο της συνάρτησης ενεργοποίησης)

Το δ για τους κρυφούς νευρώνες υπολογίζεται ως το γινόμενο
(Παράγωγος της ενεργοποίησης) x
[άθροισμα (δ*βάρος) για κάθε έξοδο του νευρώνα]



$$\delta_A(n) = \varphi'_A(v_A) \cdot [\delta_B(n) \cdot w_{AB}(n) + \delta_\Gamma(n) \cdot w_{A\Gamma}(n) + \delta_\Delta(n) \cdot w_{A\Delta}(n)]$$

$$\Delta w_{AB}(n) = \eta \cdot y_A(n) \cdot \delta_B(n)$$

$$w_{AB}(n) = w_{AB}(n) + \Delta w_{AB}(n-1)$$

δ_3	δ_4	δ_5
0.126	0.111	0.126

ΠΡΟΣ ΤΑ ΠΙΣΩ ΠΕΡΑΣΜΑ

ΝΕΥΡΩΝΑΣ 5 (νευρώνας εξόδου)

Υπολογισμός Τοπικής Κλίσης:

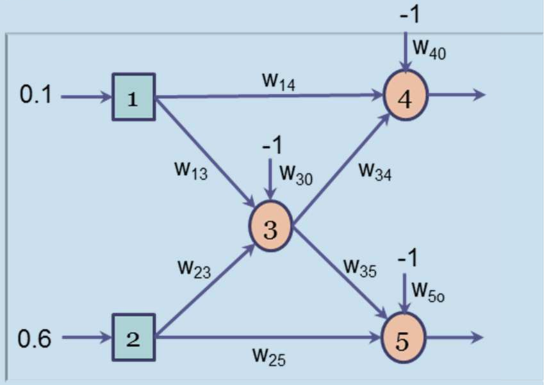
$$\delta_5 = e_5 \cdot \varphi'(v_5) = e_5 \cdot [y_5\,(1-y_5)] = 0.504 \cdot [0.496\,(1-0.496)] = 0.126$$

Διορθώσεις στα Βάρη των Ακμών:

$$\Delta w_{35} = \eta \cdot y_3 \cdot \delta_5 = 1 \cdot 0{,}473 \cdot 0{,}126 = 0{,}060$$
$$\Delta w_{25} = \eta \cdot y_2 \cdot \delta_5 = 1 \cdot 0{,}6 \cdot 0{,}126 = 0{,}076$$
$$\Delta w_{50} = \eta \cdot (-1) \cdot \delta_5 = 1 \cdot (-1) \cdot 0{,}126 = -0{,}126$$

Υπολογισμός των νέων βαρών:

$$w_{35} = w_{35} + \Delta w_{35} = 0{,}3 + 0{,}060 = 0{,}360$$
$$w_{25} = w_{25} + \Delta w_{25} = 0{,}4 + 0{,}076 = 0{,}476$$
$$w_{50} = w_{50} + \Delta w_{50} = 0{,}4 - 0{,}126 = 0{,}274$$



ΝΕΥΡΩΝΑΣ 4 (νευρώνας εξόδου)

Υπολογισμός Τοπικής Κλίσης:

…

Διορθώσεις στα Βάρη των Ακμών:

…

Υπολογισμός των νέων βαρών:

…

ΝΕΥΡΩΝΑΣ 3 (κρυφός νευρώνας)

Υπολογισμός Τοπικής Κλίσης:

$$\delta_3 = \varphi'(v_3) \cdot [w_{34} \cdot \delta_4 + w_{35} \cdot \delta_5] = y_3\,(1-y_3) \cdot [w_{34} \cdot \delta_4 + w_{35} \cdot \delta_5] = 0{,}473\,(1-0{,}473) \cdot [0{,}3 \cdot (-0{,}111) + 0{,}3 \cdot 0{,}126] = 0{,}001$$

Διορθώσεις στα Βάρη των Ακμών:

$$\Delta w_{23} = \eta \cdot y_2 \cdot \delta_3 = 1 \cdot 0{,}6 \cdot 0{,}001 = 0{,}001$$
$$\Delta w_{13} = \eta \cdot y_1 \cdot \delta_3 = 1 \cdot 0{,}1 \cdot 0{,}001 = 0$$
$$\Delta w_{30} = \eta \cdot (-1) \cdot \delta_3 = 1 \cdot (-1) \cdot 0{,}001 = -0{,}001$$

Υπολογισμός των νέων βαρών:

$$w_{23} = w_{23} + \Delta w_{23} = 0{,}4 + 0{,}001 = 0{,}401$$
$$w_{13} = w_{13} + \Delta w_{13} = 0{,}5 + 0 = 0{,}5$$
$$w_{30} = w_{30} + \Delta w_{30} = 0{,}4 - 0{,}001 = 0{,}399$$

ΠΙΝΑΚΑΣ (ΓΝΩΣΤΩΝ) ΠΑΡΑΓΩΓΩΝ:

Όνομα	Συνάρτηση	Παράγωγος
Σιγμοειδής	$\varphi(x)=\frac{1}{1+e^{-ax}}$	$\varphi'(x) = a\varphi(x)(1-\varphi(x))$
Γραμμική	$\varphi(x) = \mathrm{x}$	$\varphi'(x)=1$
Υπερβολική Εφαπτομένη	$\varphi(x) = \frac{1-e^{-ax}}{1+e^{-ax}}$	$\varphi'(x)=\frac{a}{2}[1-\varphi^2(x)]$
Γραμμική με συντελεστή	$\varphi(x) = \alpha\mathrm{x}$	$\varphi'(x)=\alpha$
Ημίτονο	$\varphi(x) = \sin(\mathrm{x})$	$\varphi'(x)=\cos(x)$
Συνημίτονο	$\varphi(x) = \cos(\mathrm{x})$	$\varphi'(x)=-\sin(x)$

ΚΡΙΤΗΡΙΟ ΤΕΡΜΑΤΙΣΜΟΥ:

- Το δίκτυο παράγει τις επιθυμητές εξόδους ή έχουν ένα σφάλμα μικρότερο από κριτήριο που έχουμε θέσει.
- Το σφάλμα παρέμεινε ίδιο σε δύο διαδοχικούς κύκλους εκπαίδευσης
- Εκτελέσαμε τον αλγόριθμο για ένα συγκεκριμένο αριθμό βημάτων.

<u>**Δομή:**</u> Ένα ΤΝΔ της δομής Hopfield δομείται ως εξής:

- N υπολογιστικούς νευρώνες και κάθε νευρώνας στέλνει την έξοδο του σε κάθε άλλο νευρώνα. Άρα έχουμε N(N-1) συνάψεις (ακμές).
- Δεν υπάρχουν αισθητήρες εισόδου ή εξόδου.

<u>**Πίνακας Βαρών:**</u> Θέλοντας να αποθηκεύσουμε τα N-διάστατα διανύσματα $X_1,X_2,X_3,\dots,X_K$ **(βασικές μνήμες)** θα χρειαστούμε ένα δίκτυο Hopfield N αισθητήρων και η αποθήκευση θα γίνει στα βάρη των ακμών με τον τύπο (Ι ο NxN μοναδιαίος πίνακας):

$$W = X_1{X_1}^T + X_2{X_2}^T + \cdots + X_K{X_K}^T - K \cdot I$$

<u>Παρατηρήσεις για τον πίνακα βαρών:</u>

- Συμμετρικός ως προς την κύρια διαγώνιο
- Η κύρια διαγώνιος είναι 0.

<u>**Έλεγχος ορθής αποθήκευσης των βασικών μνημών:**</u>

Κάθε ένα από τα Κ διανύσματα θα πρέπει να εξάγεται από τον πίνακα βαρών με την πράξη:

$sign(W \cdot X_i - \theta)$ που θα πρέπει να είναι ίσο με το X_i

<u>**Ισχύουν οι ακόλουθοι τύποι**</u> :

- Ο μέγιστος αριθμός βασικών μνημών που αποθηκεύονται χωρίς να έχουμε σφάλμα στην ανάκτηση: $M_{max} = \frac{N}{4 \cdot lnN}$
- Αντίστοιχα ώστε περισσότερες από τις μισές βασικές μνήμες να ανακαλούνται χωρίς σφάλμα: $M_{max} = \frac{N}{2 \cdot lnN}$

<u>**Πίνακας Βαρών:**</u> Θεωρείστε ότι θέλουμε να αποθηκεύσουμε τα ακόλουθα 2 διανύσματα σε ένα ΤΝΔ τύπου Hopfield 3 νευρώνων [-1,1,1], [-1,1,-1].
Το δίκτυο που αποθηκεύει τα διανύσματα πρέπει να έχει 3 αισθητήρες (αφού τα διανύσματα έχουν διάσταση 3).
Υπολογίζουμε τον πίνακα βαρών:

$$W = X_1{X_1}^T + X_2{X_2}^T - 2I$$

$$= \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & -1 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

<u>**Έλεγχος Ορθής Αποθήκευσης:**</u> Δίνεται ο πίνακας βαρών: $\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ και τα διανύσματα $X1 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$ και $X2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$. Να εξετασεί αν τα διανύσματα έχουν αποθηκευτεί σωστά στον πίνακα βαρών θεωρώντας ότι στην ανάκτηση των διανυσμάτων γίνεται χρήση του επόμενου κανόνα: «Αν η έξοδος ενός νευρώνα είναι 0, τότε θέσε την έξοδο ίση με την αντίστοιχη είσοδο»με μηδενικά κατώφλια.

Λύση:

Για το X_1:

$$sign(W \cdot X_1 - \theta) = sign\left(\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right) = sign\left(\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} \right)$$

$= sign\left(\begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} = X_1$. Αποθηκεύτηκε σωστά.

Για το X_2:

$$sign(W \cdot X_2 - \theta) = sign\left(\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right) =$$

$sign\left(\begin{bmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right) = sign\left(\begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = X_2$. Αποθηκεύτηκε σωστά.

Εισαγωγή Φθαρμένου Διανύσματος:
Εισάγεται διάνυσμα X (είτε βασική μνήμη, είτε άλλο (φθαρμένο) διάνυσμα.

Σύγχρονη Ενημέρωση Βαρών:
Γίνεται η πράξη: $sign(W \cdot X - \theta) = X'$ και
Αν το αποτέλεσμα X' = X:
- «Ισορροπεί στην X»

Αν το αποτέλεσμα X' δεν είναι ίσο με X:
- Τότε γίνεται η ίδια πράξη με είσοδο X' (οπότε ή θα έχουμε σύγκλιση ή θα έχουμε παλινδρόμηση (αποτυχία)
    - Σύγκλιση θα έχουμε όταν η είσοδος γίνει ίδια με την έξοδο. (π.χ. X=>X'=>X': ισορροπεί στην X')
    - Παλινδρόμηση θα έχουμε αν ξαναβγεί ένα διάνυσμα που είχαμε εισάγει και σε προηγούμενα (π.χ. X=>X'=>X'' και έπειτα το αποτέλεσμα X'''=X'

Ασύγχρονη Ενημέρωση Βαρών
Όπως παραπάνω, αλλά διορθώνεται μόνο μία είσοδος σε κάθε βήμα.
- Σε έναν κύκλο εκπαίδευσης θα πρέπει να διορθωθούν με τυχαία σειρά όλες οι είσοδοι.
- Μόλις ολοκληρωθεί ο κύκλος ισχύουν τα ίδια κριτήρια τερματισμού με τη σύγχρονη ενημέρωση.

Σύγχρονη Ενημέρωση Βαρών: π.χ. $W = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix}$, θ= $\begin{bmatrix} 0.5 \\ 0.5 \\ 0.5 \end{bmatrix}$

Π.χ.1: Με είσοδο το φθαρμένο διάνυσμα $\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$:

$$X(1) = sign(W \times X(0) - \theta) = sign\left(\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} - \begin{bmatrix} 0{,}5 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right)$$
$$= sign\left(\begin{bmatrix} 0 \\ -2 \\ 0 \end{bmatrix} - \begin{bmatrix} 0{,}5 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right) = sign\left(\begin{bmatrix} -0{,}5 \\ -2{,}5 \\ -0{,}5 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

Κατάσταση ισορροπίας: $[-1\ -1\ -1]^T$

Π.χ.2: Με είσοδο το φθαρμένο διάνυσμα $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$:

$$X(1) = sign(W \times X(0) - \theta) = sign\left(\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0{,}5 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$X(2) = sign(W \times X(1) - \theta) = sign\left(\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 0{,}5 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$X(3) = sign(W \times X(2) - \theta) = sign\left(\begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0{,}5 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

Παλινδρομεί μεταξύ των X(1) και X(2)

Ασύγχρονη Ενημέρωση Βαρών: (στο ίδιο δίκτυο με είσοδο $[1,1,1]^T$)
$1^{ος}$ κύκλος εκπαίδευσης:

Ο νευρώνας 1 ενημερώνεται: $sign(w_{11}x_1 + w_{12}x_2 + w_{13}x_3 - \theta_1) = sign(0 \times 1 + 1 \times 1 + (-1) \times 1 - 0{,}5) = sign(-0{,}5) = -1$ Συνεπώς: $X(1) = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

Ο νευρώνας 2 ενημερώνεται: $sign(w_{21}x_1 + w_{22}x_2 + w_{23}x_3 - \theta_2) = sign(1 \times (-1) + 0 \times 1 + 1 \times 1 - 0{,}5) = sign(-0{,}5) = -1$. Συνεπώς: $X(2) = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

Ο νευρώνας 3 ενημερώνεται: $sign(w_{31}x_1 + w_{32}x_2 + w_{33}x_3 - \theta_3) = sign((-1) \times (-1) + 1 \times (-1) + 0 \times 1 - 0{,}5) = sign(-0{,}5) = -1$. Συνεπώς: $X(3) = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$

$2^{ος}$ κύκλος εκπαίδευσης
…

Ένα ΤΝΔ της δομής Kohonen δομείται έχει **αυστηρά 2 επίπεδα:**

- Το πρώτο επίπεδο είναι το <u>επίπεδο εισόδου</u>. Αποτελείται από N νευρώνες που απλά μεταφέρουν το σήμα τους.
- Το δεύτερο επίπεδο είναι το <u>επίπεδο εξόδου (ονομάζεται και επίπεδο Kohonen)</u>. Εδώ έχουμε M νευρώνες (συνήθως σε κάποια μονοδιάστατη ή διδιάστατη δομή πλέγματος)
- Κάθε νευρώνας εισόδου <u>συνδέεται με **όλους** τους νευρώνες</u> στο επίπεδο εξόδου

Δεδομένης μιας εισόδου $[x_1, x_2, x_3, \ldots, x_N]$

- Κάθε Νευρώνας k (k=1,…M) υπολογίζει την ευκλείδια απόσταση του διανύσματος εισόδου από το διάνυσμα των βαρών των εισόδων του $[w_{k1}, w_{k2}, w_{k3}, \ldots, w_{kN}]$ σύμφωνα με τον τύπο:

$$d_k = \sqrt{\sum_{i=0}^{n} (w_{ki} - x_i)^2} = \sqrt{(w_{k1} - x_1)^2 + (w_{k2} - x_2)^2 + \cdots + (w_{kN} - x_N)^2}$$

- Ο νευρώνας που έχει την μικρότερη τιμή, είναι ο νικητής νευρώνας και παράγει έξοδο 1 (σε περίπτωση ισοπαλίας επιλέγεται τυχαία κάποιος νικητής)
- Όλοι οι υπόλοιποι νευρώνες είναι ηττημένοι νευρώνες και παράγουν έξοδο 0

Αρχικοποίηση:

Αρχικοποίηση των βαρών (από εκφώνηση ή με τυχαίο τρόπο)

a=Ρυθμός Εκπαίδευσης (από εκφώνηση)

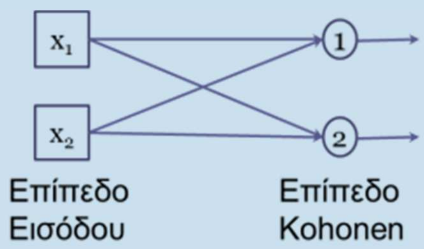
Επαναλαμβάνουμε:

- Επιλέγουμε τυχαία ένα πρότυπο και υπολογίζεται ο νικητής νευρώνας, έστω k
- Διόρθωση των βαρών **μόνο** του νικητή νευρώνα ως εξής:
- Για κάθε βάρος w_{jk}, j=1,…n (για κάθε βάρος εισόδου του νικητή)
    - Υπολογίζεται η ποσότητα: $\Delta W_{jk}=a(x_j-w_{jk})$
    - Θέτουμε $w_{jk}= w_{jk} +\Delta W_{jk}$

Έως ότου εκτελεστεί ένα πλήθος κύκλων εκπαίδευσης M

<u>Παράδειγμα:</u>

Δίνεται το δίκτυο Kohonen με 2 νευρώνες εισόδου και 2 νευρώνες εξόδου που θέλουμε να εκπαιδευτεί πάνω στα εξής πρότυπά



<u>Α=[7 8]</u>
<u>Β=[8 6]</u>
<u>Γ=[1 3]</u>
<u>Δ=[2 2]</u>
<u>Ε=[7 9]</u>
<u>Ζ=[3 3]</u>

Έστω επίσης ότι τα βάρη είναι $w_{11}=4$, $w_{12}=5$, $w_{21}=4$, $w_{22}=4$

<u>Να εκτελέσετε ένα κύκλο εκπαίδευσης χρησιμοποιώντας διαδοχικά τα πρότυπα Α,Γ,Β με α=0.5</u>

<u>Λύση:</u>

<u>Εκτελούμε με το πρότυπο Α=[7 8]. Τρέχοντα Βάρη w1=[4 4] και w2=[5 4]:</u>

1ος νευρώνας Kohonen: $d_1 = \sqrt{(w_{11} - x_1)^2 + (w_{21} - x_2)^2} = \sqrt{(4-7)^2 + (4-8)^2}=5.00$

2ος νευρώνας Kohonen: $d_2 = \sqrt{(w_{12} - x_1)^2 + (w_{22} - x_2)^2} = \sqrt{(5-7)^2 + (4-8)^2}=4.47$

Ο νικητής νευρώνας είναι ο νευρώνας 2

<u>Συνεπώς διορθώνονται τα βάρη του νευρώνα 2:</u>

$\Delta W_{12} = a(x_1 - w_{12}) = 0.5(7-5) = 1,\qquad w_{12} = w_{12} + \Delta w_{21} = 5 + 1 = 6$

$\Delta W_{22} = a(x_2 - w_{22}) = 0.5(8-4) = 2,\qquad w_{22} = w_{22} + \Delta w_{22} = 4 + 2 = 6$

- Μεταβλητός ρυθμός Εκπαίδευσης : $\alpha(n) = a(0)\left[1 - \frac{n}{N}\right]$
- Μεταβλητή Ακτίνα Νικητή Νευρώνα: $d(n) = d(0)\left[1 - \frac{n}{N}\right]$