

ΠΛΗ31

ΕΝΟΤΗΤΑ 3: ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Μάθημα 3.2: Δίκτυα ενός Νευρώνα - Εκπαίδευση

Δημήτρης Ψούνης



ΠΕΡΙΕΧΟΜΕΝΑ

A.Θεωρία

1. Εκπαίδευση ΤΝΔ

1. Εισαγωγή
2. Μάθηση με Επίβλεψη
3. Μάθηση χωρίς Επίβλεψη
4. Μάθηση με Ενίσχυση

2. Εκπαίδευση ενός νευρώνα

1. Ο κανόνας μάθησης Δέλτα
2. Ο αλγόριθμος του κανόνα μάθησης δέλτα
3. Παράδειγμα εκπαίδευσης με τον κανόνα μάθησης δέλτα
4. Παρατηρήσεις

B.Ασκήσεις

1. Εφαρμογές

A. Θεωρία

1. Εκπαίδευση ΤΝΔ

1. Εισαγωγή

Η εκπαίδευση ενός ΤΝΔ είναι μία υπολογιστική διαδικασία στην οποία «εκπαιδεύουμε το ΤΝΔ προκειμένου να κάνει την επιθυμητή εργασία, δηλαδή:

- Να απαντά ΝΑΙ, για κάθε θετικό πρότυπο.
- Να απαντά ΌΧΙ για κάθε αρνητικό πρότυπο.

Σε ένα ΤΝΔ ενός νευρώνα, η εκπαίδευση συνίσταται στο να υπολογιστούν:

- Τα βάρη των εισόδων w_1, w_2, \dots, w_k
- Η τιμή του κατωφλίου θ .

Ήδη έχουμε μάθει μία διαδικασία εκπαίδευσης ενός ΤΝΔ:

- Η εύρεση των βαρών και του κατωφλίου με την διαδικασία του γραφικού διαχωρισμού των προτύπων σε θετικά και αρνητικά στιγμιότυπα.

Γενικά οι διαδικασίες εκπαίδευσης ενός ΤΝΔ χωρίζονται σε 3 κατηγορίες:

- Την μάθηση με επίβλεψη (ΕΝΤΟΣ ύλης)
- Την μάθηση χωρίς επίβλεψη (ΕΝΤΟΣ ύλης)
- Την μάθηση με ενίσχυση (εκτός ύλης)

A. Θεωρία

1. Εκπαίδευση ΤΝΔ

2. Μάθηση με Επίβλεψη

Στην μάθηση με επίβλεψη:

- Υπάρχουν τα πρότυπα ως ζεύγη (είσοδος – επιθυμητή έξοδος)

Επαναλαμβάνεται μια αλγοριθμική διαδικασία στην οποία:

- Δίδονται εισόδοι στο ΤΝΔ
- Το ΤΝΔ δίνει την απάντησή του.
- Συγκρίνεται η απάντηση του ΤΝΔ με την επιθυμητή έξοδο
- Γίνεται προσαρμογή των παραμέτρων (κατώφλια – βάρη) από το ίδιο το σύστημα.

Το σύστημα προσπαθεί να ανακαλύψει τη συσχέτιση μεταξύ των δεδομένων εισόδου και της επιθυμητής εξόδου προσαρμόζοντας τις συναρτήσεις που προσομοιώνουν την διαδικασία που υπολογίζεται.

A. Θεωρία

1. Εκπαίδευση ΤΝΔ

3. Μάθηση χωρίς Επίβλεψη

Στην μάθηση χωρίς επίβλεψη:

- Τα πρότυπα δίδονται μόνο ως εισόδοι (δεν δίνονται στο σύστημα οι επιθυμητές εξόδοι)

Με αλγοριθμικές διαδικασίες ζητείται από το σύστημα

- Να μάθει πως σχετίζονται τα δεδομένα εισόδου.
- Να βρεί μόνο του τις απαντήσεις που πρέπει να δίδονται.

Το σκεπτικό είναι ότι δεδομένα που «μοιάζουν» αρκετά μεταξύ τους, θα έχουν την ίδια απάντηση. Οπότε το σύστημα προσπαθεί να ομαδοποιήσει τα δεδομένα εισόδου σε διαφορετικές ομάδες.

A. Θεωρία

1. Εκπαίδευση ΤΝΔ

4. Μάθηση με Ενίσχυση

Στην μάθηση με ενίσχυση:

- Παραλλαγή της μάθησης με επίβλεψη στην οποία παρέχεται πληροφορία στο ΤΝΔ και από τον εξωτερικό επιβλέποντα.

Συγκεκριμένα

- Ο επιβλέπων ενημερώνει το σύστημα για το αν έδωσε σωστή ή λάθος απάντηση.
- Ο επιβλέπων δεν δίνει την σωστή απάντηση, αλλά με απόρριψη της απάντησης της προτεινόμενης εξόδου, οδηγεί το σύστημα στο να ανακαλύψει την σωστή απάντηση.

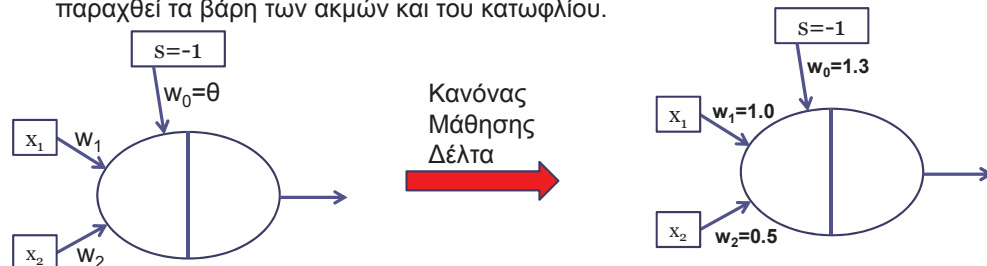
A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

1. Ο κανόνας μάθησης Δέλτα

Ο κανόνας μάθησης Δέλτα είναι ένας αλγόριθμος που χρησιμοποιείται για την εκπαίδευση ενός νευρώνα

- Μας παρέχονται τα πρότυπα εισόδου-εξόδου, τρέχουμε τον αλγόριθμο και έχουν παραχθεί τα βάρη των ακμών και του κατωφλίου.



| x1 | x2 | Y |
|----|----|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

2. Ο αλγόριθμος του κανόνα μάθησης Δέλτα

Ο κανόνας μάθησης Δέλτα χρησιμοποιείται για την εκπαίδευση ενός Perceptron

Αρχικοποίηση:

- Αρχικοποιούμε τα διανύσματα:
 - $W = [w_0, w_1, \dots, w_n]$ με τα αρχικοποιημένα βάρη των ακμών
 - Για κάθε πρότυπο $1, \dots, K$: 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 : $0 < \eta < 1$
- Κύκλος Εκπαίδευσης:** Για κάθε πρότυπο: $i = 1 \dots K$:
 - Υπολόγισε το δυναμικό για το πρότυπο i ως: $v = W^T \cdot x_i$
 - Υπολόγισε την έξοδο από την συνάρτηση δυναμικού: $y_i = \phi(v)$
 - Υπολόγισε το σφάλμα ως: $e = d_i - y_i$
 - Αν το σφάλμα δεν είναι μηδενικό
 - Υπολογίζονται νέα βάρη ως: $W_{new} = W_{old} + \eta * e * x_i$

Με την ολοκλήρωση ενός κύκλου εκπαίδευσης: Γίνεται έλεγχος τερματισμού του αλγορίθμου. Συνήθως ελέγχονται τα εξής κριτήρια:

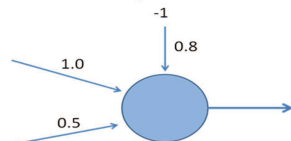
- Ο νευρώνας απάντησε σωστά σε όλα τα πρότυπα στον κύκλο εκπαίδευσης
- Ο νευρώνας δεν βελτίωσε καθόλου την απόδοσή του.
- Ο αλγόριθμος πέρασε ένα άνω όριο των βημάτων επανάληψης

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

3. Παράδειγμα Εκπαίδευσης με τον Κανόνα Δέλτα

Θέλουμε να εκπαιδεύσουμε έναν αισθητήρα ώστε να επιλύει το πρόβλημα του λογικού ΚΑΙ. Κατά τη διάρκεια της διαδικασίας εκπαίδευσης με ρυθμό εκπαίδευσης $\eta=0.5$, έχουμε τα βάρη που φαίνονται στο Σχήμα 2. Ο αισθητήρας ακολουθεί το μοντέλο McCulloch-Pitts 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:

Κωδικοποίηση 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: $\eta = 0.5$

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

3. Παράδειγμα Εκπαίδευσης με τον Κανόνα Δέλτα

1ος Κύκλος Εκπαίδευσης:

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

$$W^T \cdot x_1 = [0.8, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -0.8$$

• Αφού $-0.8 < 0$ η έξοδος είναι $y_1 = 0$.

• Σφάλμα: $\text{error} = d_1 - y_1 = 0$. Τα βάρη δεν αλλάζουν.

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

$$W^T \cdot x_2 = [0.8, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = -0.3$$

• Αφού $-0.3 < 0$ η έξοδος είναι $y_2 = 0$.

• Σφάλμα: $\text{error} = d_2 - y_2 = 0$. Τα βάρη δεν αλλάζουν.

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

$$W^T \cdot x_3 = [0.8, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = -0.8 + 1.0 = 0.2$$

• Αφού $0.2 \geq 0$ η έξοδος είναι $y_3 = 1$.

• Σφάλμα: $\text{error} = d_3 - y_3 = -1$. Τα βάρη αλλάζουν

$$W = W + \eta \cdot \text{error} \cdot x_3 = \begin{bmatrix} 0.8 \\ 1.0 \\ 0.5 \end{bmatrix} + 0.5 * (-1) * \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.0 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0.5 \\ -0.5 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.3 \\ 0.5 \\ 0.5 \end{bmatrix}$$

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

$$W^T \cdot x_4 = [1.3, 0.5, 0.5] \times \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = -1.3 + 0.5 + 0.5 = -0.3$$

• Αφού $-0.3 < 0$ η έξοδος είναι $y_4 = 0$.

• Σφάλμα: $\text{error} = d_4 - y_4 = 1$. Τα βάρη αλλάζουν

$$W = W + \eta \cdot \text{error} \cdot x_3 = \begin{bmatrix} 1.3 \\ 0.5 \\ 0.5 \end{bmatrix} + 0.5 * 1 * \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.3 \\ 0.5 \\ 0.5 \end{bmatrix} + \begin{bmatrix} -0.5 \\ 0.5 \\ 0.5 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.0 \\ 1.0 \end{bmatrix}$$

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

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

3. Παράδειγμα Εκπαίδευσης με τον Κανόνα Δέλτα

2ος Κύκλος Εκπαίδευσης:

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

$$W^T \cdot x_1 = [0.8, 1.0, 1.0] \times \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -0.8$$

• Αφού $-0.8 < 0$ η έξοδος είναι $y_1 = 0$.

• Σφάλμα: $\text{error} = d_1 - y_1 = 0$. Τα βάρη δεν αλλάζουν.

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

$$W^T \cdot x_2 = [0.8, 1.0, 1.0] \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = -0.8 + 1.0 = 0.2$$

• Αφού $0.2 \geq 0$ η έξοδος είναι $y_2 = 1$.

• Σφάλμα: $\text{error} = d_2 - y_2 = -1$. Τα βάρη αλλάζουν.

$$W = W + \eta \cdot \text{error} \cdot x_3 = \begin{bmatrix} 0.8 \\ 1.0 \\ 1.0 \end{bmatrix} + 0.5 * (-1) * \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 1.0 \\ 1.0 \end{bmatrix} + \begin{bmatrix} 0.5 \\ 0 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 1.3 \\ 1.0 \\ 0.5 \end{bmatrix}$$

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

$$W^T \cdot x_3 = [1.3, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = -1.3 + 1.0 = -0.3$$

• Αφού $-0.3 < 0$ η έξοδος είναι $y_3 = 0$.

• Σφάλμα: $\text{error} = d_3 - y_3 = 0$. Τα βάρη δεν αλλάζουν

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

$$W^T \cdot x_4 = [1.3, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = -1.3 + 1 + 0.5 = 0.2$$

• Αφού $0.2 \geq 0$ η έξοδος είναι $y_4 = 1$.

• Σφάλμα: $\text{error} = d_4 - y_4 = 0$ Τα βάρη δεν αλλάζουν

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

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

3. Παράδειγμα Εκπαίδευσης με τον Κανόνα Δέλτα

3ος Κύκλος Εκπαίδευσης:

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

$$W^T \cdot x_1 = [1.3, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} = -1.3$$

• Αφού $-1.3 < 0$ η έξοδος είναι $y_1 = 0$.

• Σφάλμα: $\text{error} = d_1 - y_1 = 0$. Τα βάρη δεν αλλάζουν.

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

$$W^T \cdot x_2 = [1.3, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = -1.3 + 0.5 = -0.8$$

• Αφού $-0.8 < 0$ η έξοδος είναι $y_2 = 0$.

• Σφάλμα: $\text{error} = d_2 - y_2 = 0$. Τα βάρη δεν αλλάζουν.

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

$$W^T \cdot x_3 = [1.3, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = -1.3 + 1.0 = -0.3$$

• Αφού $-0.3 < 0$ η έξοδος είναι $y_3 = 0$.

• Σφάλμα: $\text{error} = d_3 - y_3 = 0$. Τα βάρη δεν αλλάζουν

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

$$W^T \cdot x_4 = [1.3, 1.0, 0.5] \times \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = -1.3 + 1 + 0.5 = 0.2$$

• Αφού $0.2 \geq 0$ η έξοδος είναι $y_4 = 1$.

• Σφάλμα: $\text{error} = d_4 - y_4 = 0$ Τα βάρη δεν αλλάζουν

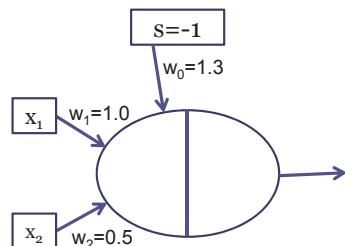
Ο νευρώνας απάντησε σωστά σε όλα τα πρότυπα. Ο αλγόριθμος τερματίζει.

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

3. Παράδειγμα Εκπαίδευσης με τον Κανόνα Δέλτα

Συνεπώς ο νευρώνας που κατασκευάστηκε και αναγνωρίζει την λογική πύλη AND είναι ο εξής:



A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

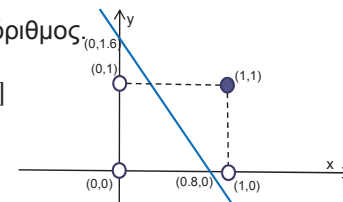
3. Παράδειγμα Εκπαίδευσης με τον Κανόνα Δέλτα

Μελετούμε τις διαδοχικές ευθείες που παράγαγε ο αλγόριθμος

Τα αρχικά βάρη που μας είχαν δοθεί $W = [0.8, 1.0, 0.5]$

Αντιστοιχούν στην συνάρτηση: $x + 0.5y - 0.8 = 0$

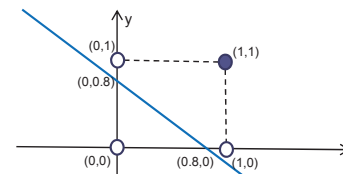
Περνά από τα σημεία: (0,1.6) και (0.8,0)



Μετά τον 1^ο κύκλο εκπαίδευσης: $W = [0.8, 1.0, 1.0]$

Αντιστοιχούν στην συνάρτηση: $x + y - 0.8 = 0$

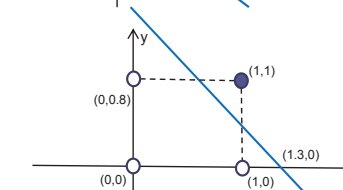
Περνά από τα σημεία: (0,0.8) και (0.8,0)



Μετά τον 2^ο κύκλο εκπαίδευσης: $W = [1.3, 1.0, 0.5]$

Αντιστοιχούν στην συνάρτηση: $x + 0.5y - 1.3 = 0$

Περνά από τα σημεία: (0,2.6) και (1.3,0)



Ο 3^{ος} κύκλος εκπαίδευσης: Επιβεβαιώνει τα βάρη.

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

4. Παρατηρήσεις για τον κανόνα Δέλτα

- Αν το πρόβλημα είναι γραμμικά διαχωρίσιμο, τότε ο κανόνας μάθησης Δέλτα, θα συγκλίνει (Η απόδειξη είναι εκτός ύλης)
- Αν το πρόβλημα δεν είναι διαχωρίσιμο, τότε ο αλγόριθμος θα συνεχίζει να μετατοπίζει την ευθεία επ' άπειρον και γι' αυτό απαιτείται κριτήριο τερματισμού μετά από κάποιο πλήθος επαναλήψεων.
- Υπάρχει η δυνατότητα να χρησιμοποιήσουμε διαφορετική συνάρτηση ενεργοποίησης και οσοδήποτε εισόδους.
- Το πλήθος των επαναλήψεων εξαρτάται από την παράμετρο μάθησης η : Μικρό η οδηγεί σε μεγάλο αριθμό επαναλήψεων και μεγάλο η οδηγεί σε μικρό πλήθος επαναλήψεων.
 - Όσο πιο πολύ μεγάλη τιμή του η οδηγεί σε βιαίες μεταβολές της εξίσωσης ευθείας με αποτέλεσμα να μην μπορεί να σταθεροποιηθεί ώστε να διαχωρίζει τα δεδομένα

A. Θεωρία

2. Εκπαίδευση Νευρώνα με τον κανόνα μάθησης Δέλτα

4. Παρατηρήσεις για τον κανόνα Δέλτα

- Οι κύκλοι εκπαίδευσης αναφέρονται και ως εποχές εκπαίδευσης. Ενδέχεται το διάνυσμα των βαρών να μας δίδεται και με μία παράμετρο που συμβολίζει την χρονική στιγμή. Έτσι αν π.χ. έχουμε ένα σύνολο 4 προτύπων τότε οι κύκλοι εκπαίδευσης του W μπορούν να συμβολιστούν και ως εξής:
 - 1^{ος} κύκλος εκπαίδευσης: $W(0), W(1), W(2), W(3)$
 - 2^{ος} κύκλος εκπαίδευσης: $W(4), W(5), W(6), W(7)$
 - 3^{ος} κύκλος εκπαίδευσης: $W(8), W(9), W(10), W(11)$
 -
- Όπου συμβολίζουμε διαφορετικά τα βάρη σε κάθε μελέτη προτύπου που κάνουμε, δηλαδή π.χ.
 - αντί να γράφουμε: $W_{new} = W_{old} + \eta * e * x_i$
 - θα γράφουμε: $W(t) = W(t-1) + \eta * e * x_i$

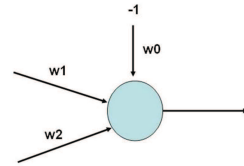


Β. Ασκήσεις

Εφαρμογή 1

Θεωρήστε έναν αισθητήρα με δύο εισόδους που είναι στη διαδικασία εκπαίδευσης της συναρτήσεως του λογικού “OR” και έχει συνάρτηση ενεργοποίησης που δίνεται από την σχέση

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



Κατά τη διαδικασία εκπαίδευσης, στον κύκλο n τα βάρη έχουν αποκτήσει τις τιμές $w_0 = \theta = 0$, $w_1 = 0.5$ και $w_2 = 0.5$. Συνεχίστε την διαδικασία εκπαίδευσης ώστε να εκπαιδευτεί ο αισθητήρας στην επίλυση του λογικού “OR”, δεδομένου ότι η σταθερά εκπαίδευσης είναι $\eta = 0.5$.