Ασκήσεις Φροντιστηρίου «Υπολογιστική Νοημοσύνη Ι»

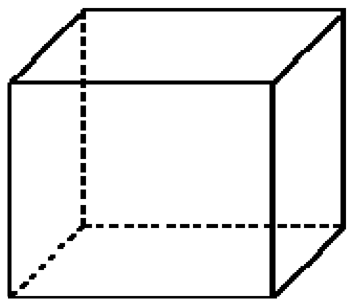
5ο Φροντιστήριο

Πρόβλημα 1ο

Δίνεται το παρακάτω σύνολο εκπαίδευσης:

#	**Είσοδος**	**Κατηγορία**
1	0 0 0	Α
2	1 0 0	Α
3	1 1 0	Β
4	0 1 0	Α
5	0 1 1	Β
6	0 0 1	Α
7	1 0 1	Β
8	1 1 1	Β

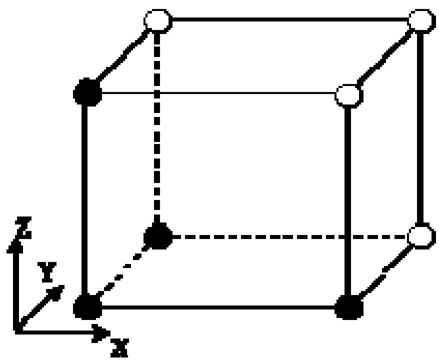
α) Στον παρακάτω κύβο τοποθετείστε τα οκτώ παραδείγματα, χρησιμοποιώντας διαφορετικά σύμβολα για να διαχωρίσετε τα σημεία της κατηγορίας Α από αυτά της κατηγορίας Β. Ορίστε το σύστημα συντεταγμένων που χρησιμοποιείτε (αρχή, κατεύθυνση και μονάδες των αξόνων). Ελέγξτε εάν τα σημεία των δύο κατηγοριών είναι γραμμικώς διαχωρίσιμα. **(1)**



β) Κατασκευάστε ένα αισθητήρα (perceptron), ο οποίος διαχωρίζει τις δύο κατηγορίες παραδειγμάτων. **(1.5)**

Απάντηση:

α) Ονομάζοντας τις τρεις εισόδους ως x, y, z αντίστοιχα, ορίζουμε τρεις άξονες που ξεκινούν από την κάτω αριστερή γωνία του κύβου και με θετικές κατευθύνσεις δεξιά, μέσα και επάνω αντίστοιχα. Θεωρούμε ότι η πλευρά του κύβου ισούται με τη μονάδα. Χρησιμοποιούμε έναν σκιασμένο κύκλο για να συμβολίσουμε τα σημεία της κατηγορίας A και έναν μη-σκιασμένο κύκλο για να συμβολίσουμε τα σημεία της κατηγορίας B. Όλα τα παραπάνω φαίνονται στο σχήμα που ακολουθεί:



Από το παραπάνω σχήμα φαίνεται ότι οι δύο κατηγορίες είναι γραμμικά διαχωρίσιμες.

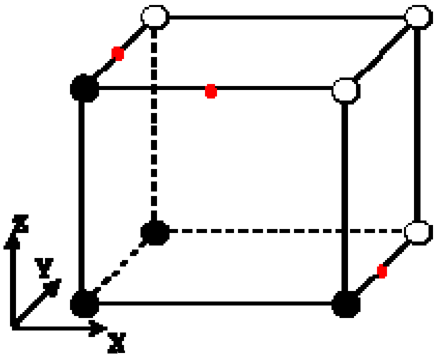
β) Ο αισθητήρας θα έχει τρεις εισόδους, συν την τάση πόλωσης, άρα συνολικά χρειαζόμαστε τέσσερα βάρη. Δεν θα πραγματοποιήσουμε εκπαίδευση, αλλά θα βρούμε τα βάρη απευθείας από τη γεωμετρία του προβλήματος. Ειδικότερα, θα βρούμε ένα επίπεδο το οποίο να διαχωρίζει τα 8 σημεία. Η εξίσωση ενός επιπέδου στο χώρο των τριών διαστάσεων είναι της μορφής ax+by+cz+d=0. Οι συντελεστές a, b και c θα είναι τα βάρη των εισόδων x, y και z αντίστοιχα, ενώ, αν θεωρήσουμε τάση πόλωσης ίση με 1, η σταθερά d θα είναι το βάρος της τάσης πόλωσης (ουσιαστικά η σταθερά d ισούται με το γινόμενο της τάσης πόλωσης επί το βάρος της τάσης πόλωσης).

Για να υπολογίσουμε την εξίσωση ενός επιπέδου χρειαζόμαστε 3 σημεία. Αυτά μπορούν να βρίσκονται στα μέσα των ακμών που συνδέουν ένα σημείο της κατηγορίας A με ένα σημείο της κατηγορίας B. Τέτοια είναι τα:

0.5, 0, 1

0, 0.5, 1

1, 0.5, 0



Έχουμε λοιπόν τις παρακάτω εξισώσεις:

$$0.5a+c+d=0$$

$$0.5b+c+d=0$$

$$a+0.5b+d=0$$

Αφαιρώντας τη δεύτερη από την τρίτη εξίσωση παίρνουμε:

$$a-c=0 \Leftrightarrow a=c$$

Οπότε από την πρώτη εξίσωση έχουμε:

$$1.5a+d=0 \text{ ή } a=-d/1.5$$

Το σύστημα έχει άπειρες λύσεις (τέσσερις άγνωστοι για τρεις εξισώσεις), οπότε μπορούμε να δώσουμε αυθαίρετες μη-μηδενικές τιμές σε μία από τις παραμέτρους. Έστω d=-3, για να πάρουμε απλά νούμερα για τις παραμέτρους a και c.

Έχουμε λοιπόν:

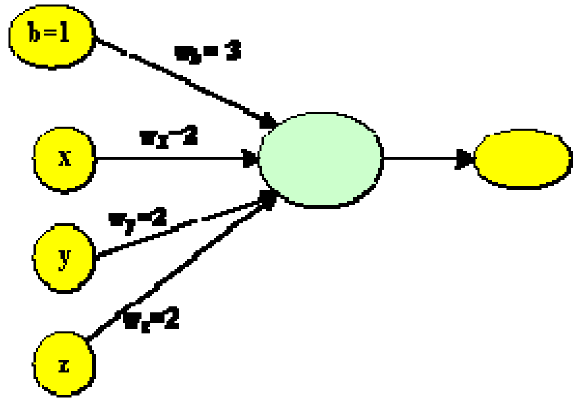
$$a=c=2,\ d=-3$$

και από τη δεύτερη εξίσωση έχουμε b=2.

Άρα η εξίσωση του επιπέδου είναι η:

$2x+2y+2z-3=0$

Ο αισθητήρας λοιπόν είναι ο παρακάτω:

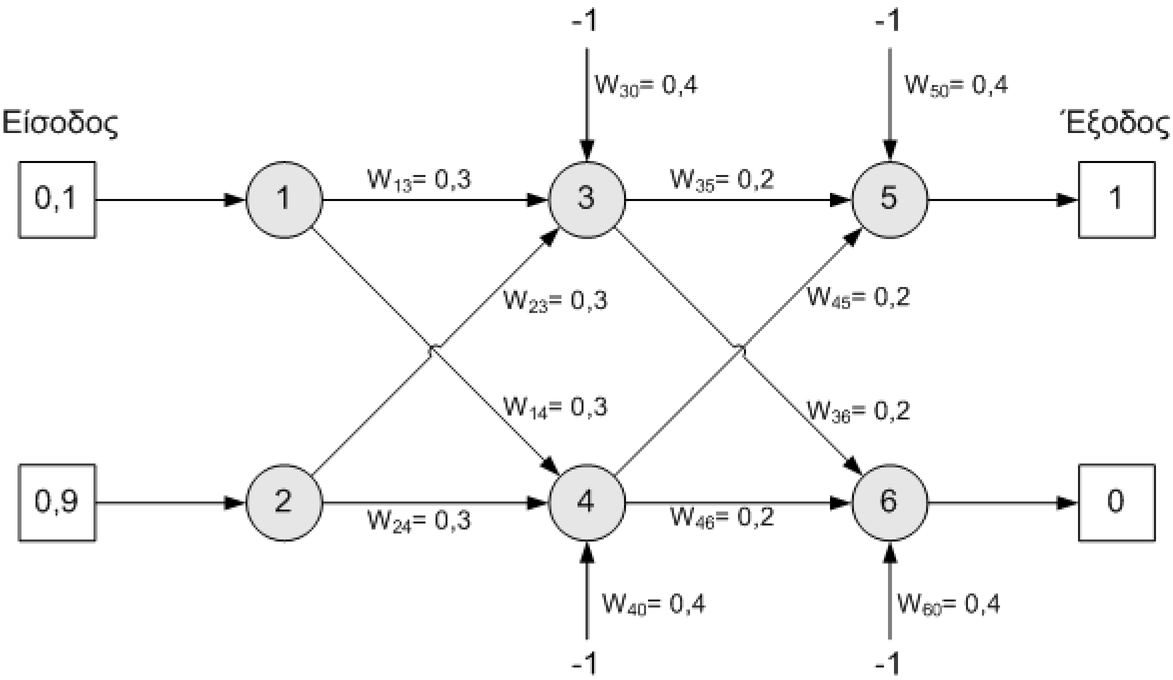


Ένας τελευταίος έλεγχος που έχουμε να κάνουμε, είναι να ελέγξουμε την έξοδο του αισθητήρα για τα διάφορα παραδείγματα. Βλέπουμε λοιπόν ότι για όλα τα παραδείγματα της κατηγορίας Α η έξοδος του αισθητήρα είναι 0 ενώ για όλα τα παραδείγματα της κατηγορίας Β η έξοδος του αισθητήρα είναι 1.

Παλιά Θέματα

Θέμα 3^{ο} (Σεπτέμβριος 2005)

Έστω ότι έχουμε ένα Νευρωνικό Δίκτυο εμπρόσθιας τροφοδότησης με αρχιτεκτονική 2-2-2 και την αρχικοποίηση των βαρών όπως φαίνεται στο παρακάτω σχήμα.



Για τους νευρώνες του κρυφού επιπέδου και του επιπέδου εξόδου χρησιμοποιούμε σα συνάρτηση ενεργοποίησης τη λογιστική ($\varphi(\upsilon) = \frac{1}{1+e^{-\upsilon}}$). Χρησιμοποιώντας παράμετρο μάθησης η = 0.25 υλοποιείστε μια επανάληψη του αλγορίθμου Πίσω-Διάδοσης του Λάθους.

Σημείωση: Οι υπολογισμοί θα γίνουν με ακρίβεια 3 δεκαδικών ψηφίων. Ως βοήθημα στις αριθμητικές πράξεις δίνονται τα παρακάτω:

$$\frac{1}{1+e^{0.1}} = 0.475 \quad \textit{και} \quad \frac{1}{1+e^{0.21}} = 0.448 \text{ .}$$

ΑΠΑΝΤΗΣΗ

<u>Υπολογισμός της εξόδου με τα αρχικά βάρη.</u>

Άθροισμα στο κόμβο 3: $U_3 = (w_{13}*0.1) + (w_{23}*0.9) – 0.4 = 0.3*0.1+0.3*0.9 -0.4 = -0.1$

Ενεργοποίηση στον κόμβο 3: $F_3 = F(-0.1) = \frac{1}{1+e^{0.1}} = 0.475$

Άθροισμα στο κόμβο 4: $U_4 = (w_{14}*0.1)+ (w_{24}*0.9) – 0.4 = 0.3*0.1+0.3*0.9 – 0.4 = -0.1$

Ενεργοποίηση στον κόμβο 4: $F_4 = F(-0.1) = 0.475$

Άθροισμα στο κόμβο 5: $U_5 = (w_{35}* F_3) + (w_{45}* F_4) – 0.4 = 0.2*0.475+0.2*0.475 - 0.4 = -0.21$

Ενεργοποίηση στον κόμβο 5: $F_5 = F(-0.21) = \frac{1}{1+e^{0.1}} = 0.448$

Άθροισμα στο κόμβο 6: $U_6 = (w_{36}* F_3)+ (w_{46}* F_4) – 0.4 = 0.2*0.475+0.2*0.475 – 0.4 = - 0.21$

Ενεργοποίηση στον κόμβο 6: F_6=F(-0.21) = 0.448

Άρα η έξοδος του ΤΝΔ με τα δοθέντα βάρη είναι (0.448,0.448).

<u>Ανανέωση των βαρών</u>

Ανανέωση των συνδέσεων μεταξύ κρυφού και επιπέδου εξόδου

Η έξοδος του ΤΝΔ έπρεπε να ήταν [1.0,0.0] για το πρότυπο εισόδου [0.1,0.9]. Άρα το σφάλμα στην έξοδο είναι:

Για τον κόμβο 5: (1-0.448)

Για τον κόμβο 6: (0-0.448)

Ο υπολογισμός των βαρών ακολουθεί την παρακάτω διαδικασία.

Για τους νευρώνες στο επίπεδο εξόδου ισχύει:

$$\delta_{pj} = (t_{pj} - o_{pj}) * f_j'(S_{sj})$$

Για το πρόβλημά μας έχουμε:

$$\delta_{pj} = (t_{pj} - o_{pj}) * o_{pj}(1 - o_{pj})).$$

Άρα:

$\delta_5 = (t - o) * o * (1 - o)$, όπου t = target =1.0 και o=output=0.448.

Συνεπώς δ_5=0.137.

Ομοίως: $\delta_6 = (t - o) * o * (1 - o)$, όπου t = target =0.0 και o=output=0.448.

Συνεπώς δ_6=-0.111.

Για τον υπολογισμό των νέων τιμών w_{45}, w_{35}, w_{36} και w_{46} ισχύουν:

$$\begin{aligned}
w_{35_new} &= w_{35} + n * (F_3) * \delta_5 = 0.2 + 0.25 * 0.475 * 0.137 = 0.216 \\
w_{45_new} &= w_{45} + n * (F_4) * \delta_5 = 0.2 + 0.25 * 0.475 * 0.137 = 0.216 \\
w_{36_new} &= w_{36} + n * (F_3) * \delta_6 = 0.2 + 0.25 * 0.475 * (-0.111) = 0.187 \\
w_{46_new} &= w_{46} + n * (F_4) * \delta_6 = 0.2 + 0.25 * 0.475 * (-0.111) = 0.187
\end{aligned}$$

Όπου: n=0.25 είναι ο ρυθμός εκμάθησης

F_3 είναι η ενεργοποίηση της σιγμοειδούς στον κόμβο 3 και

F_4 είναι η ενεργοποίηση της σιγμοειδούς στον κόμβο 4.

Για τις νέες τιμές των κατωφλίων ισχύει:

$$\begin{aligned}
w_{50_new} &= w_{50} + n * (y_{50}) * \delta_5 = 0.4 + 0.25 * (-1) * 0.137 = 0.366 \\
w_{60_new} &= w_{60} + n * (y_{60}) * \delta_6 = 0.4 + 0.25 * (-1) * (-0.111) = 0.428
\end{aligned}$$

<u>Ανανέωση των συνδέσεων μεταξύ κρυφού και επιπέδου εισόδου</u>

Ο υπολογισμός των βαρών ακολουθεί την παρακάτω διαδικασία.

Για τους νευρώνες στο κρυφό επίπεδο ισχύει:

$$\delta_{pj} = f_j'(S_{sj}) \sum_k \delta_{pk} * w_{kj}$$

Επειδή εδώ έχουμε δύο κόμβους εξόδου, σε κάθε νευρώνα του κρυφού επιπέδου έχουμε δύο συνδέσεις με το επίπεδο εξόδου. Άρα στη παραπάνω σχέση πρέπει να προσέξουμε το άθροισμα $\sum_k \delta_{pk} * w_{kj}$ (σε αυτό το σημείο διαφοροποιείται η παρούσα άσκηση με αυτές που είδαμε ως τώρα).

Οπότε:

$\delta_3 = f_j'(S_{sj})\sum_k \delta_{pk} * w_{kj}$ (ΠΡΟΣΟΧΗ: ΕΧΟΥΜΕ ΑΘΡΟΙΣΜΑ ΓΙΝΟΜΕΝΩΝ!!!)

$\delta_3 = F_3*(1-F_3)*[(\delta_5*w_{35}) + (\delta_6*w_{36})]$ = 0.475*(1-0.475)* [(0.137*0.2) + (-0.111*0.2)] = 0.001

Ομοίως:

$\delta_4=F_4*(1-F_4)*$ $[(\delta_5*w_{45})+(\delta_6*w_{46})]$ = 0.475*(1-0.475)*[(0.137*0.2) + (-0.111*0.2)] =0.001

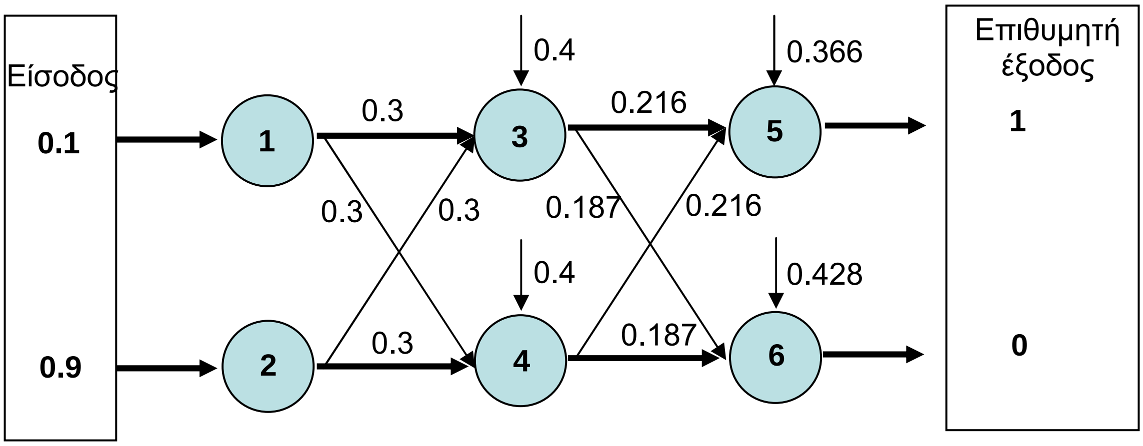
Για τον υπολογισμό των νέων τιμών w_{13} , w_{14}, w_{23} και w_{24} ισχύουν:

$$w_{13\_new} = w_{13} + n*(i_1)*\delta_3 = 0.3 + 0.25*0.1*(0.001) \approx 0.3$$
$$w_{14\_new} = w_{14} + n*(i_1)*\delta_4 = 0.3 + 0.25*0.1*(0.003) \approx 0.3$$
$$w_{23\_new} = w_{23} + n*(i_2)*\delta_3 = 0.3 + 0.25*0.9*(0.001) \approx 0.3$$
$$w_{24\_new} = w_{24} + n*(i_2)*\delta_4 = 0.3 + 0.25*0.9*(0.001) \approx 0.3$$
$$w_{30\_new} = w_{30} + n*(y_{30})*\delta_3 = 0.4 + 0.25*(-1)*(0.001) \approx 0.4$$
$$w_{40\_new} = w_{40} + n*(y_{40})*\delta_4 = 0.4 + 0.25*(-1)*(0.001) \approx 0.4$$

Όπου: n=0.25 είναι ο ρυθμός εκμάθησης, i_1=0.1 και i_2=0.9 είναι οι είσοδοι του ΤΝΔ.

Όλες οι μεταβολές είναι πολύ μικρές και δεν μπορούν να φανούν με ακρίβεια 3 δεκαδικών ψηφίων. Άρα οι τιμές των βαρών και των κατωφλίων θεωρούμε ότι δεν μεταβάλλονται.

Το ΤΝΔ λοιπόν μετά από ένα κύκλο εκπαίδευσης και τη μεταβολή των βαρών είναι:



Θέμα 1^o (Σεπτέμβριος 2005)

Α) Η λογική συνάρτηση NAND από *n* δυαδικές εισόδους $x_1, x_2, x_3 \ldots x_n$ ορίζεται από την εξίσωση $\neg(x_1 \wedge x_2 \wedge x_3 \wedge \ldots \wedge x_n)$. Έστω ότι το Αληθές αναπαριστάται από την τιμή 1 και το Ψευδές από την τιμή 0 και $n \geq 2$ ένας γνωστός ακέραιος αριθμός. Μπορεί αυτή η συνάρτηση να αναπαρασταθεί από ένα νευρώνα Perceptron? Εάν ναι κατασκευάστε έναν τέτοιο Perceptron, αλλιώς δικαιολογείστε την απάντηση σας.

ΑΠΑΝΤΗΣΗ

ΝΑΙ. Είναι γραμμικά διαχωρίσιμο το πρόβλημα, γιατί θα πρέπει να βγάζει 0 όταν όλες οι είσοδοι είναι 1.

Ένας Perceptron που μπορεί να το κάνει αυτό έχει n εισόδους και μια έξοδο. Όλα τα βάρη είναι -1 και το κατώφλι θα είναι *(–n + 0,5)*, Οπότε η έξοδος θα δίνεται από τον τύπο

$$y = \sum_{1}^{n} (-1) * x_i - (-n + 0{,}5) = \sum_{1}^{n} (-1) * x_i + n - 0{,}5)$$

Β) Η λογική συνάρτηση SAME δυο δυαδικών εισόδων x_1 και x_2 παίρνει τιμή 1 όταν και οι δυο είσοδοι έχουν την ίδια τιμή (και οι δύο να είναι 1 ή και οι δυο να είναι 0). Μπορεί αυτή η συνάρτηση να αναπαρασταθεί από ένα νευρώνα Perceptron? Εάν ναι κατασκευάστε έναν τέτοιο Perceptron, αλλιώς δικαιολογείστε την απάντηση σας.

ΑΠΑΝΤΗΣΗ

ΟΧΙ. Γιατί δεν είναι γραμμικά διαχωρίσιμη συνάρτηση.

Θέμα 4^o (Φεβρουάριος 2005)

Θέλουμε να υλοποιήσουμε ένα Νευρωνικό δίκτυο το οποίο θα λειτουργεί ως κωδικοποιητής-αποκωδικοποιητής με την έννοια ότι μπορεί να κωδικοποιεί πρότυπα εισόδου μήκους N bits σε ένα πρότυπο μήκους $\log_2(N)$ και στη συνέχεια αποκωδικοποιεί αυτή την αναπαράσταση στο πρότυπο εξόδου μήκους N bits. Θεωρούμε ότι το δίκτυο έχει ένα κρυφό επίπεδο και τα πρότυπα εισόδου και εξόδου είναι όπως τα παρακάτω:

Πρότυπα Εισόδου	**Πρότυπα Εξόδου**
10000000	10000000
01000000	01000000
00100000	00100000
00010000	00010000
00001000	00001000
00000100	00000100
00000010	00000010
00000001	00000001

Να απαντήσετε στα παρακάτω ερωτήματα

(α) (10%) Περιγράψτε και σχεδιάστε την αρχιτεκτονική του νευρωνικού δικτύου.

(β) (5%) Πόσοι συνολικά νευρώνες απαιτούνται για την υλοποίησή του.

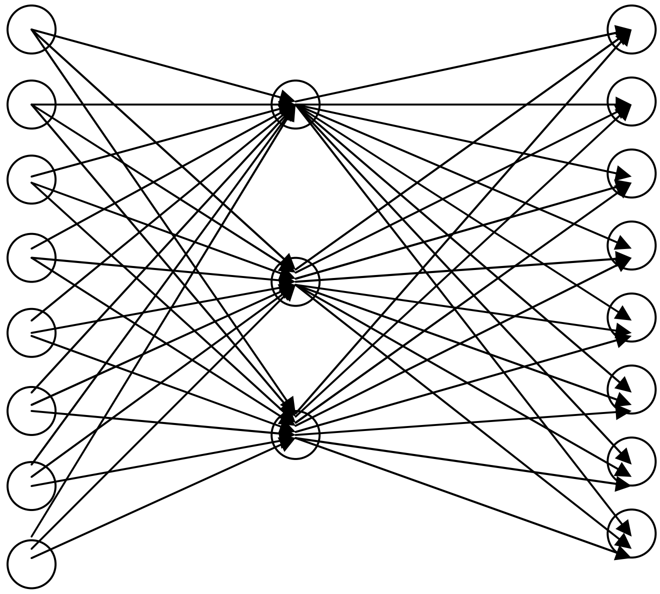
(γ) (5%) Εξετάστε αν κρίνετε απαραίτητη τη χρήση κατωφλίου ενεργοποίησης των νευρώνων.

(δ) (5%) Προτείνετε την καταλληλότερη, κατά τη γνώμη σας, συνάρτηση ενεργοποίησης.

(ε) (5%) Ποιες είναι οι διαστάσεις του πίνακα βαρών (συμπεριλαμβανομένων των κατωφλίων ενεργοποίησης αν κρίνετε ότι χρειάζονται)

ΑΠΑΝΤΗΣΗ

(α) Το ζητούμενο δίκτυο θα αποτελείται από τρία επίπεδα. Το επίπεδο εισόδου με N=8 νευρώνες, ένα κρυφό επίπεδο με 3 νευρώνες (γιατί $\log_2 8 = 3$) και το επίπεδο εξόδου με επίσης 8 νευρώνες. Το N είναι ίσο με 8 γιατί τα δοσμένα πρότυπα εκπαίδευσης (τόσο για την είσοδο, όσο και για την έξοδο) αποτελούνται από 8 δυαδικά ψηφία.



Επίπεδο Εισόδου **Κρυφό επίπεδο** **Επίπεδο εξόδου**

8 Νευρώνες **3 Νευρώνες** **8 Νευρώνες**

Σε σχέση με τη συνδεσμολογία των νευρώνων είναι λογικό να υιοθετήσουμε την γενική περίπτωση (όλα-με-όλα), δηλαδή όλοι οι νευρώνες ενός επίπέδου συνδέονται με όλους τους νευρώνες του επόμενου επιπέδου. Η αρχιτεκτονική αυτή φαίνεται στο παραπάνω σχήμα

(β) Όπως φαίνεται στο Σχήμα του ερωτήματος (α) απαιτούνται συνολικά 8+3+8 = 19 νευρώνες.

(γ) Λόγω της φύσης του προβλήματος είναι απαραίτητη η χρήση κατωφλίου ενεργοποίησης των νευρώνων.

(δ) Η φύση των δεδομένων στα πρότυπα εκπαίδευσης (δυαδικά ψηφία) συνηγορεί για τη χρήση της συνάρτησης κατωφλίου σα συνάρτηση ενεργοποίησης. Ωστόσο, επειδή ο καταλληλότερος αλγόριθμος εκπαίδευσης του δικτύου είναι αυτός της Πίσω Διάδοσης του λάθους, η προτεινόμενη συνάρτηση ενεργοποίησης θα πρέπει να είναι η λογιστική συνάρτηση *1/(1+exp(-αu))*, (όπου *α* η παράμετρος κλήσης). Άλλωστε, για μεγάλες τιμές της παραμέτρου *α* η σιγμοειδής συνάρτηση συμπεριφέρεται σα συνάρτηση κατωφλίου.

(ε) Με βάση την αρχιτεκτονική του προτεινόμενου δικτύου που φαίνεται στο Σχήμα του ερωτήματος (α), οι πίνακες συνδέσεων έχουν διαστάσεις 9x3 (επίπεδο εισόδου προς το κρυφό επίπεδο) και 3x9 (κρυφό επίπεδο προς το επίπεδο εξόδου) με την σημείωση ότι το κατώφλι ενεργοποίησης των νευρώνων έχει συμπεριληφθεί ως συνδετική παράμετρος.

<u>ΘΕΜΑ 1^o</u> (2.5 μονάδες)

Σχεδιάστε έναν αισθητήρα (perceptron) δύο εισόδων (βηματική συνάρτηση ενεργοποίησης), ο οποίος υλοποιεί τη λογική συνάρτηση A v ¬B.

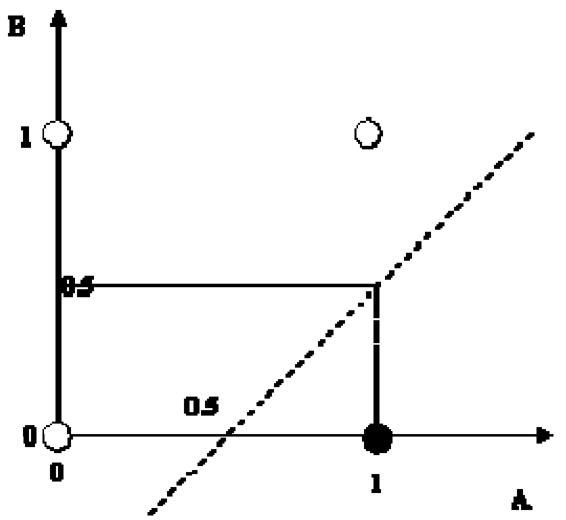
<u>Υπόδειξη</u>: Μην εκτελέσετε εκπαίδευση με τον κανόνα δέλτα, αλλά βρείτε μια ευθεία η οποία διαχωρίζει τα παραδείγματα εκπαίδευσης, υπολογίστε τις παραμέτρους της και από αυτές υπολογίστε τα βάρη του αισθητήρα.

Απάντηση:

Έστω ότι έχουμε τέσσερα παραδείγματα, για όλους τους συνδυασμούς δυαδικών τιμών για τις μεταβλητές εισόδου A και B:

#	A	B	Έξοδος (A v ¬B)
1	0	0	0
2	0	1	0
3	1	0	1
4	1	1	0

Σε δισδιάστατο διάγραμμα σχεδιάζουμε μια ευθεία η οποία διαχωρίζει τα τέσσερα παραδείγματα:



Στο παραπάνω διάγραμμα με λευκούς κύκλους φαίνονται τα παραδείγματα που αντιστοιχούν στην έξοδο 0 και με μαύρο κύκλο το παράδειγμα που αντιστοιχεί στην έξοδο 1. Τα παραδείγματα είναι φανερά γραμμικώς διαχωρίσιμα. Η διακεκομμένη ευθεία κλίσης 45^o διαχωρίζει σωστά τα παραδείγματα (και μάλιστα το κάνει με τον καλύτερο δυνατό τρόπο).

Θα προσπαθήσουμε να βρούμε την εξίσωση αυτής της ευθείας. Αυτή θα είναι της μορφής α*A+β*B+γ=0, όπου αναζητάμε τους συντελεστές α, β και γ.

Η ευθεία αυτή διέρχεται από δύο γνωστά σημεία, τα (0.5, 0) και (1,0.5). Άρα η εξίσωσή της ικανοποιεί τις σχέσεις:

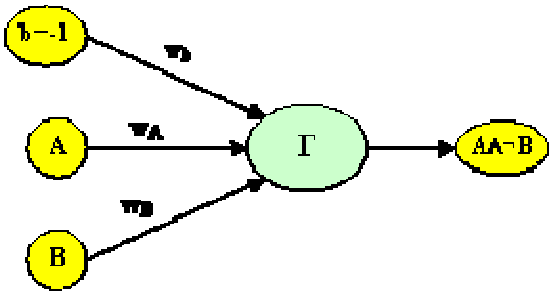
$$\alpha * 0.5 + \gamma = 0$$

$$\alpha * 1 + \beta * 0.5 + \gamma = 0$$

Θέτοντας αυθαίρετα την τιμή γ=1 παίρνουμε α=-2 και β=2. Άρα η εξίσωση της ευθείας γίνεται:

$$-2*A+2*B+1=0 \quad (1)$$

Παρακάτω φαίνεται το μοντέλο του αισθητήρα. Θεωρούμε τάση πόλωσης ίση με -1.



Ο αισθητήρας παράγει έξοδο 1 όταν ισχύει:

$$w_A*A+w_B*B-w_\Gamma \geq 0 \quad (2)$$

Συγκρίνοντας την εξίσωση (2) με την (1) προκύπτει ότι αυτές ταυτίζονται όταν w_A=-2, w_B=2 και w_Γ=1.

Στο σημείο αυτό χρειάζεται προσοχή, γιατί η εξίσωση (1) θα μπορούσε να γραφεί και ως:

$$2*A-2*B-1=0 \quad (3)$$

όπου η (3) προέκυψε από την (1) με πολλαπλασιασμό των δύο μερών της με -1. Οι εξισώσεις (1) και (3) διαφέρουν κατά το ποιο ημιεπίπεδο από τα δύο στα οποία χωρίζουν το επίπεδο αντιστοιχεί σε θετικές τιμές της παράστασης στο αριστερό σκέλος τους.

Αντικαθιστώντας στο αριστερό σκέλος της (1) τις συντεταγμένες (A=1, B=0) παίρνουμε:

$$-2*1+2*1+1=1>0$$

Άρα πράγματι η (1) είναι αυτή που πρέπει να λάβουμε υπόψη για τον υπολογισμό των βαρών του αισθητήρα, όπως και κάναμε παραπάνω.