

Κεφάλαιο 19

Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Nets)

Τεχνητή Νοημοσύνη - Β' Έκδοση

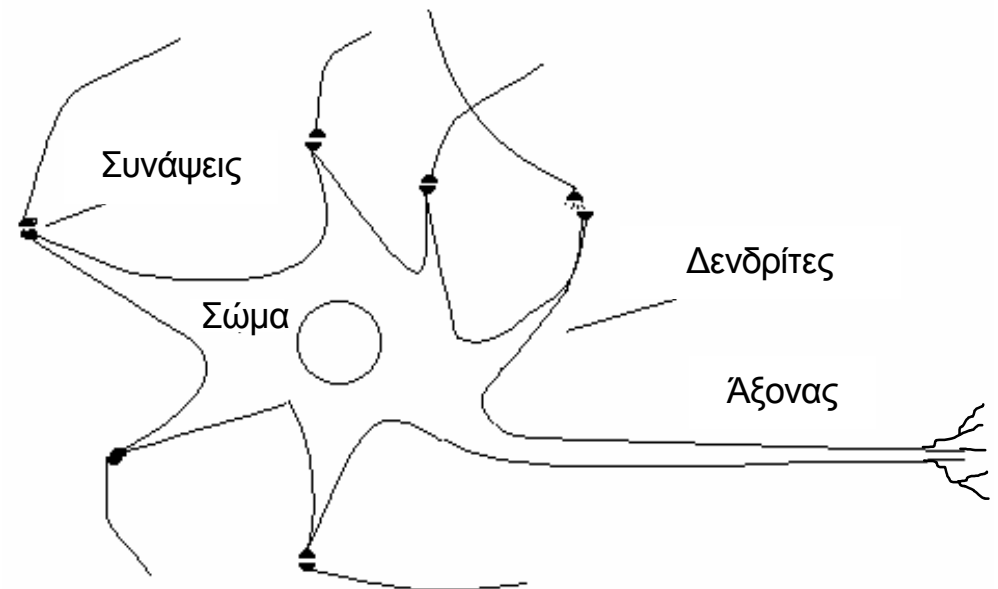
Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Η. Σακελλαρίου

Νευρωνικά Δίκτυα (ΝΔ) - Εισαγωγή

- ❖ Είναι μια ιδιαίτερη προσέγγιση στη δημιουργία συστημάτων με νοημοσύνη.
 - ❑ δεν αναπαριστούν ρητά τη γνώση
 - ❑ δεν υιοθετούν ειδικά σχεδιασμένους αλγόριθμους αναζήτησης.
- ❖ Βασίζονται σε βιολογικά πρότυπα (ανθρώπινο εγκέφαλο)

Βιολογικός Νευρώνας

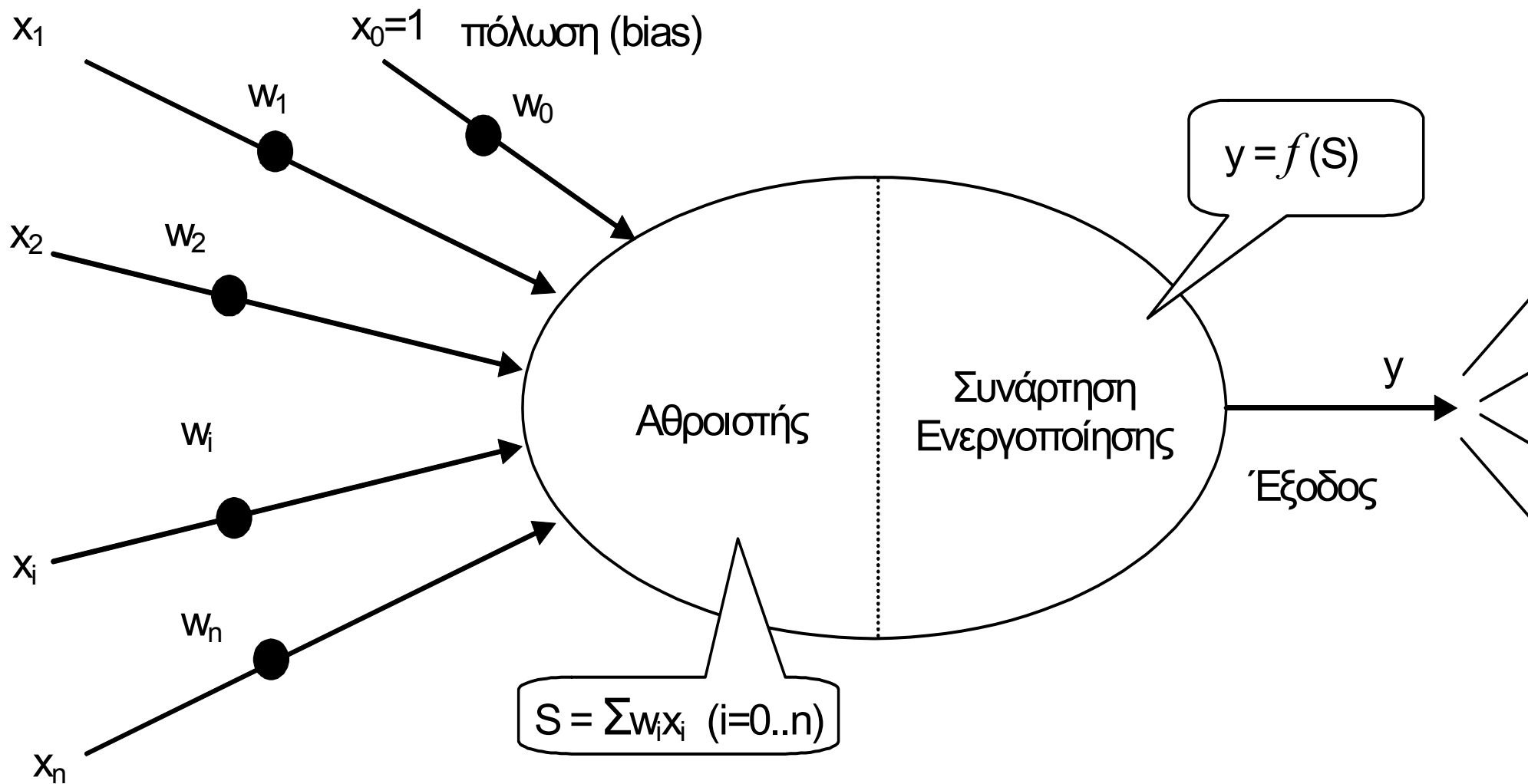
- ❖ μάθηση και μνήμη: μεταβολή στην αγωγιμότητα των συνάψεων
- ❖ Τα σήματα που εισέρχονται στο *σώμα* μέσω των *δενδριτών*, συνδυάζονται και αν το αποτέλεσμα ξεπερνά κάποιο *κατώφλι*, διαδίδεται μέσω του *άξονα* προς άλλους νευρώνες.



Φυσικά Νευρωνικά Δίκτυα

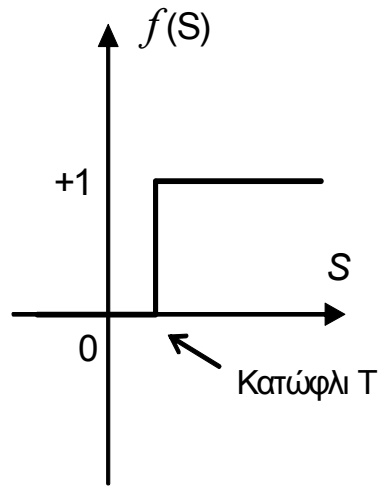
- ❖ ανθρώπινος εγκέφαλος:
 - ❑ ~100 δισεκατομμύρια νευρώνες .
 - ❑ κάθε νευρώνας συνδέεται κατά μέσο όρο με 1000 άλλους νευρώνες
 - ❑ ~100 τρισεκατομμύρια συνάψεις
- ❖ η αντιγραφή είναι εφικτή μόνο περιορισμένη κλίμακα
- ❖ χρόνος απόκρισης των βιολογικών νευρώνων: της τάξης msec
 - ❑ ...αλλά, λαμβάνει πολύπλοκες αποφάσεις, εκπληκτικά γρήγορα.
- ❖ η υπολογιστική ικανότητα του εγκεφάλου και η πληροφορία που περιέχει είναι διαμοιρασμένα σε όλο του τον όγκο
 - ❑ παράλληλο και κατανεμημένο υπολογιστικό σύστημα.

Μοντέλο Τεχνητού Νευρώνα (*artificial neuron*)

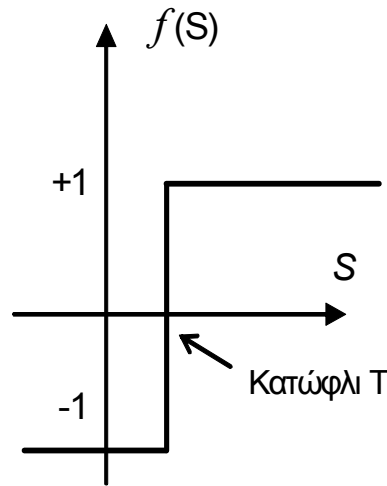


Συναρτήσεις Ενεργοποίησης

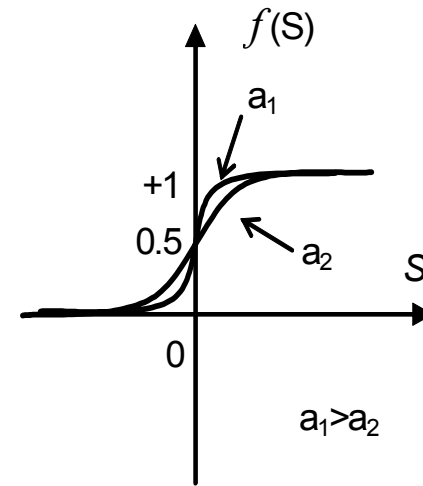
- ❖ Βασική απαίτηση: να είναι μη γραμμική ώστε να μπορεί να μοντελοποιεί μη γραμμικά φαινόμενα.



α) Βηματική Συνάρτηση



β) Συνάρτηση Προσήμου

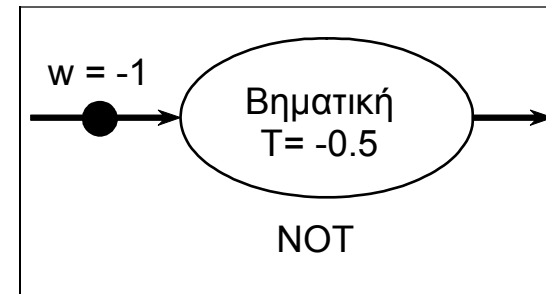
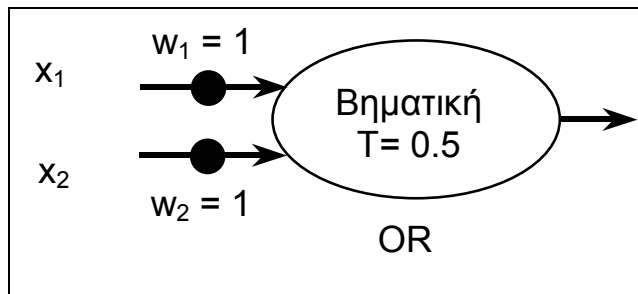
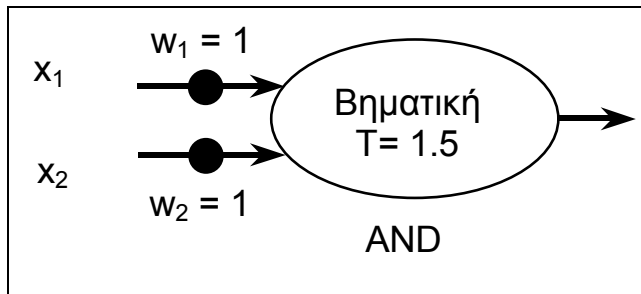


γ) Λογιστική Συνάρτηση

- ❖ Η *λογιστική* (*logistic*) συνάρτηση - μέλος οικογένειας *σιγμοειδών* συναρτήσεων.

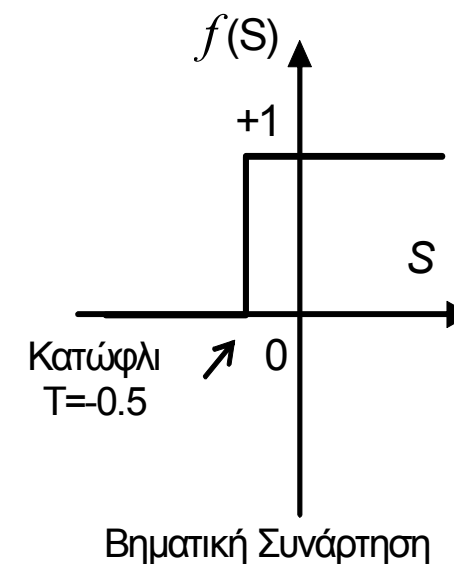
$$\Phi(S) = \frac{1}{1 + e^{-a \cdot S}}$$

Υλοποίηση Λογικών Συναρτήσεων με Τεχνητό Νευρώνα



❖ **Παράδειγμα:** υλοποίηση του NOT:

- βηματική συνάρτηση ενεργοποίησης με κατώφλι $T = -0.5$

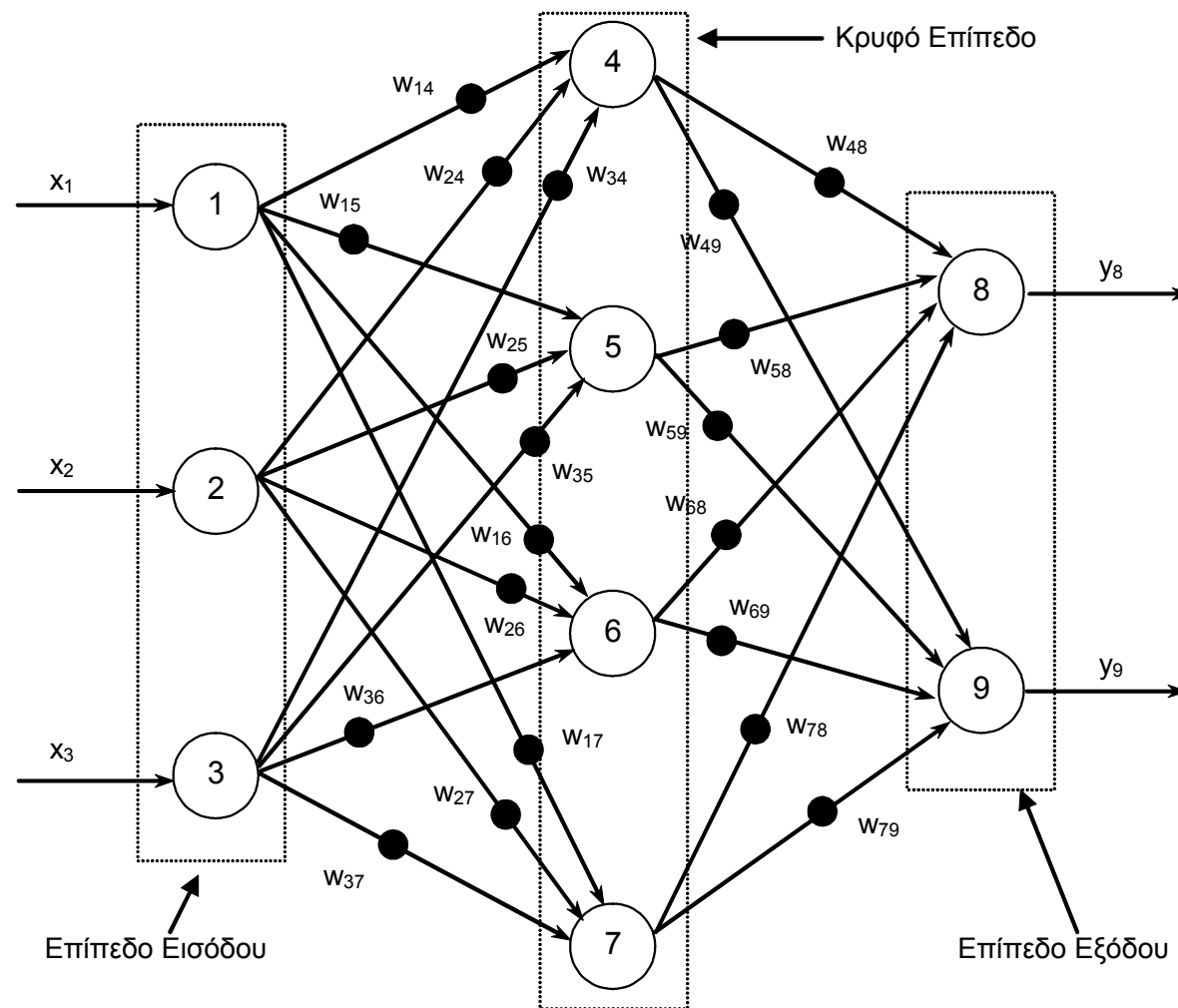


Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)

❖ Συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος τεχνητών νευρώνων οργανωμένων σε δομές παρόμοιες με αυτές του ανθρώπινου εγκεφάλου.

❖ Συντομογραφία για πολυεπίπεδα ΤΝΔ: $(p, m_1, m_2, \dots, m_q, n)$

Πλήρως
συνδεδεμένο
ΤΝΔ απλής
τροφοδότησης
3-4-2.



Χαρακτηριστικά – Ορολογία

- ❖ Οι νευρώνες των διαφόρων στρωμάτων μπορεί να είναι:
 - ❑ Πλήρως συνδεδεμένοι (*fully connected*)
 - ❑ Μερικώς συνδεδεμένοι (*partially connected*)

- ❖ Τα ΤΝΔ χαρακτηρίζονται ως:
 - ❑ Δίκτυα με πρόσθια τροφοδότηση (*feedforward*)
 - ❑ Δίκτυα με ανατροφοδότηση (*feedback ή recurrent*)

- ❖ Στην πλειοψηφία των εφαρμογών χρησιμοποιούνται δίκτυα απλής τροφοδότησης.

Μάθηση και Ανάκληση

- ❖ **Μάθηση** - *learning* (ή *εκπαίδευση* - *training*) είναι η διαδικασία της τροποποίησης της τιμής των βαρών του δικτύου, ώστε δοθέντος συγκεκριμένου διανύσματος εισόδου να παραχθεί συγκεκριμένο διάνυσμα εξόδου.
- ❖ **Ανάκληση** (*recall*) είναι η διαδικασία του υπολογισμού ενός διανύσματος εξόδου για συγκεκριμένο διάνυσμα εισόδου και τιμές βαρών.
- ❖ 3 είδη μάθησης:
 - ❑ **Μάθηση υπό Επίβλεψη** (*supervised learning*)
 - ❑ **Μάθηση χωρίς Επίβλεψη** (*unsupervised learning*)
 - ❑ **Βαθμολογημένη Μάθηση** (*graded learning*)

Στην πράξη, στις περισσότερες εφαρμογές ΤΝΔ χρησιμοποιείται μάθηση υπό επίβλεψη, για την οποία υπάρχουν αρκετοί αλγόριθμοι.

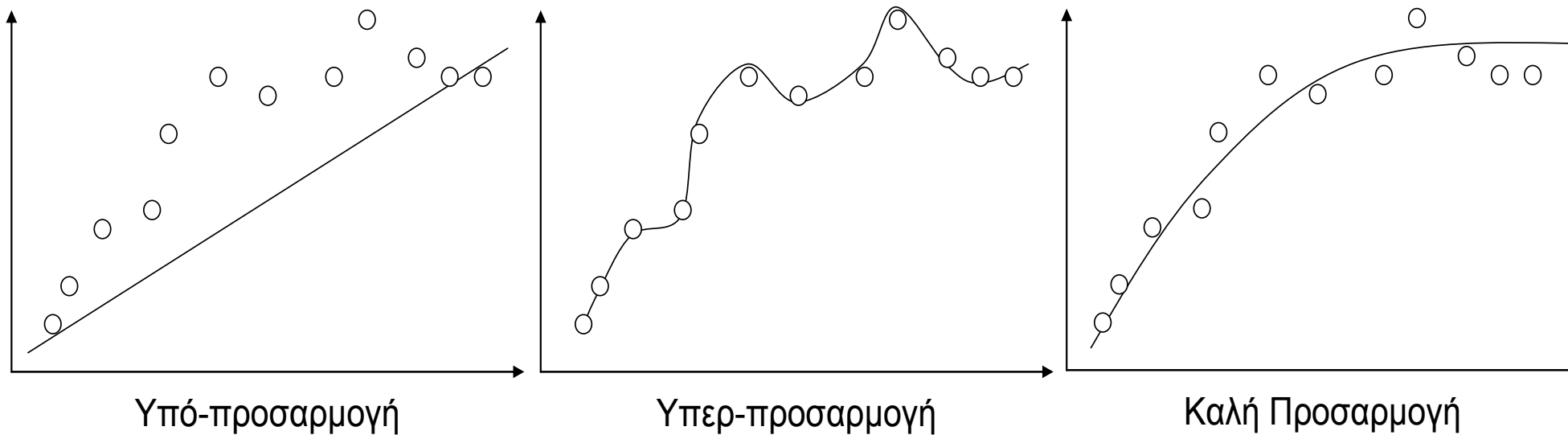


Αλγόριθμοι Μάθησης υπό Επίβλεψη

- ❖ *Κανόνας Δέλτα (Delta rule learning)*
- ❖ *Αλγόριθμος ανάστροφης μετάδοσης λάθους (back propagation)*
- ❖ *Ανταγωνιστική μάθηση (competitive learning)*
- ❖ *Τυχαία μάθηση (random learning)*

Χαρακτηριστικά Εκπαίδευσης

- ❖ υποπροσαρμογής ή ατελούς μάθησης (*underfitting*)
- ❖ υπερπροσαρμογής (*overfitting*)



Δεδομένα Εκπαίδευσης

- ❖ χρήση σε κύκλους εκπαίδευσης που ονομάζονται *εποχές* (*epochs*)
 - ❑ μάθηση δέσμης (*batch learning*)
 - ❑ επαυξητική μάθηση (*incremental learning*)
 - ❑ συνδυασμός των δύο παραπάνω μεθόδων
- ❖ Η εκπαίδευση τερματίζεται όταν το κριτήριο ελέγχου της ποιότητας του δικτύου φτάσει σε κάποια επιθυμητή τιμή.
- ❖ Κριτήρια Ελέγχου Ποιότητας
 - ❑ μέσο σφάλμα του συνόλου εκπαίδευσης
 - ❑ μεταβολή του μέσου σφάλματος του συνόλου εκπαίδευσης
- ❖ Κανονικοποίηση δεδομένων εκπαίδευσης και ελέγχου
 - ❑ (τα δεύτερα, με βάση τις παραμέτρους κανονικοποίησης των πρώτων).



Βασικές Ιδιότητες των ΤΝΔ

- ❖ Ικανότητα να μαθαίνουν μέσω παραδειγμάτων (learn by example).
- ❖ Η δυνατότητα θεώρησής τους ως κατανεμημένη μνήμη (distributed memory) και ως μνήμη συσχέτισης (associative memory).
- ❖ Η μεγάλη τους ανοχή σε σφάλματα (fault-tolerant).
- ❖ Η εξαιρετική ικανότητά τους για αναγνώριση προτύπων (pattern recognition).

ΤΝΔ Πρόσθιας Τροφοδότησης (*feedforward*)

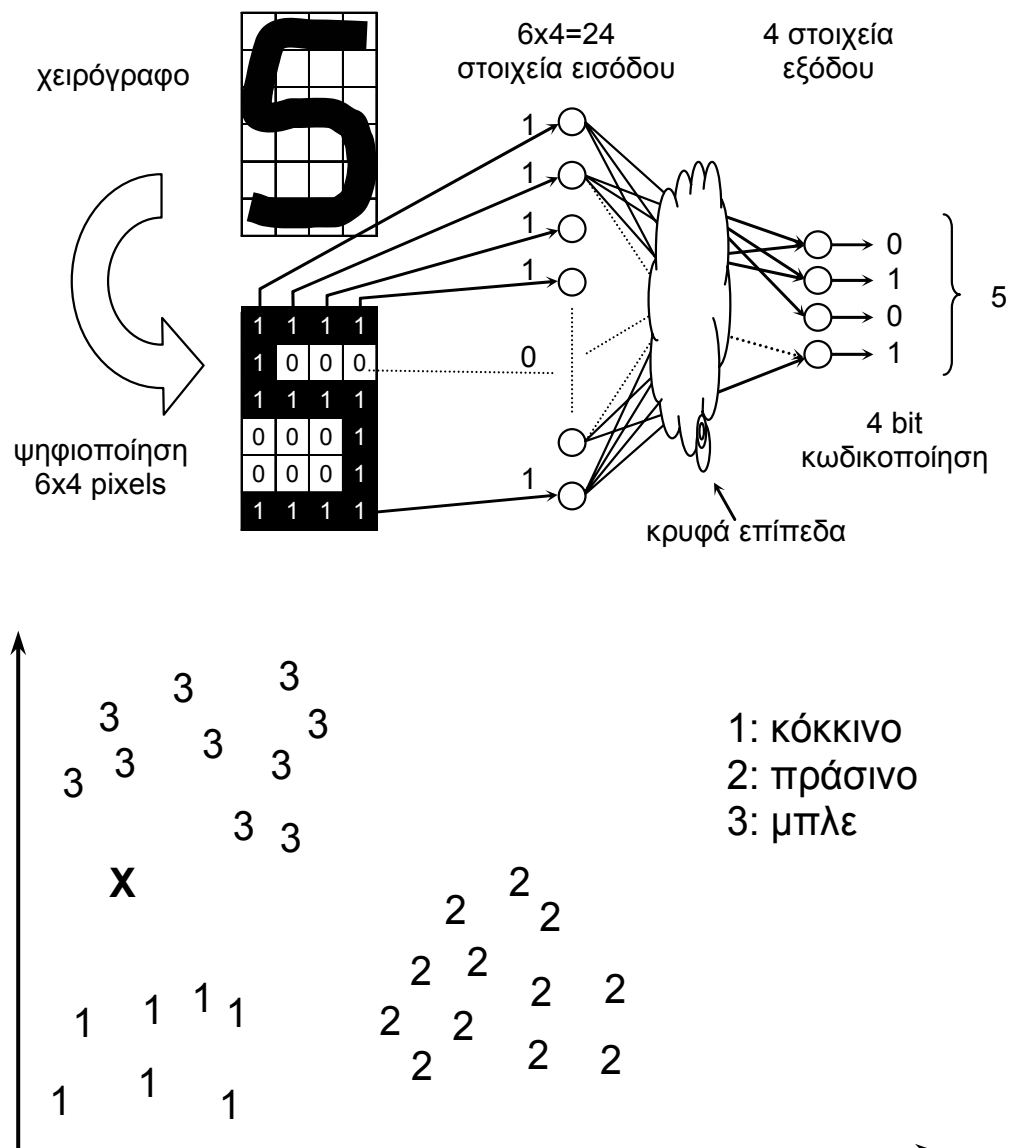
❖ επίπεδο εισόδου, επίπεδο εξόδου, κανένα, ένα ή περισσότερα κρυφά επίπεδα

❖ **Είδος μάθησης:** μάθηση με επίβλεψη.

❖ **Τοπολογία** του δικτύου

❑ δεν υπάρχει κανόνας για τον προσδιορισμό κρυφών επιπέδων, νευρώνων ανά επίπεδο, συνδεσμολογίας

❑ Τα δεδομένα εισόδου-εξόδου βοηθούν στην εκτίμηση του αριθμού νευρώνων στα επίπεδα εισόδου και εξόδου. π.χ.:



Κρυφά Επίπεδα

- ❖ Ο αριθμός των νευρώνων στα κρυφά επίπεδα σχετίζεται με πολύπλοκο τρόπο με:
 - ❑ τον αριθμό των νευρώνων στα επίπεδα εισόδου και εξόδου,
 - ❑ τον αριθμό των διανυσμάτων εκπαίδευσης και την ύπαρξη ή όχι θορύβου σε αυτά,
 - ❑ την πολυπλοκότητα της συνάρτησης ή της κατηγοριοποίησης που πρέπει να μάθει το ΤΝΔ
 - ❑ τις συναρτήσεις ενεργοποίησης που χρησιμοποιούνται,
 - ❑ τον αλγόριθμο εκπαίδευσης, κτλ.

- ❖ Εμπειρικός κανόνας για προβλήματα κατηγοριοποίησης:
 - ❑ αριθμός νευρώνων στα κρυφά επίπεδα < αριθμό διανυσμάτων εκπαίδευσης
 - ❑ **αιτία:** για να αποφευχθεί απομνημόνευση

- ❖ Συνήθως κάθε νευρώνας συνδέεται με όλους τους νευρώνες του επόμενου επιπέδου.

- ❖ Απαιτούνται αρκετές δοκιμές και πειραματισμοί.

Perceptron

- ❖ Η πιο απλή τοπολογία δικτύου με απλή τροφοδότηση.
 - ❑ ένας νευρώνας, βηματική συνάρτηση, μάθηση με επίβλεψη
- ❖ αλγόριθμος μεταβολής βαρών

Μέχρις ότου ικανοποιηθεί η συνθήκη τερματισμού της εκπαίδευσης επανέλαβε:
Για κάθε ζευγάρι εισόδου x και επιθυμητής εξόδου t από το σύνολο εκπαίδευσης

1. Υπολόγισε την έξοδο y
2. Εάν $y=t$ τότε δε γίνεται καμία μεταβολή στα βάρη
3. Εάν $y \neq t$ τότε μετέβαλε τα βάρη των ενεργών γραμμών εισόδου (αυτών που έχουν σήμα $\neq 0$) κατά την ποσότητα $\Delta w = d \cdot (t - y) \cdot x$ έτσι ώστε το y να πλησιάσει το t .

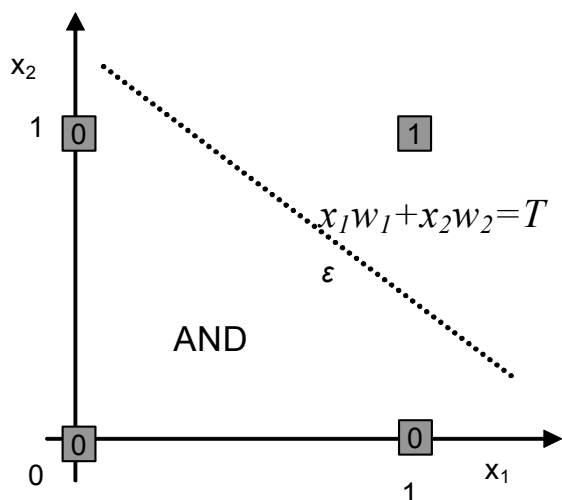
- ❖ d : ρυθμός μάθησης (*learning rate*)

Γραμμική Διαχωρισιμότητα (1/2)

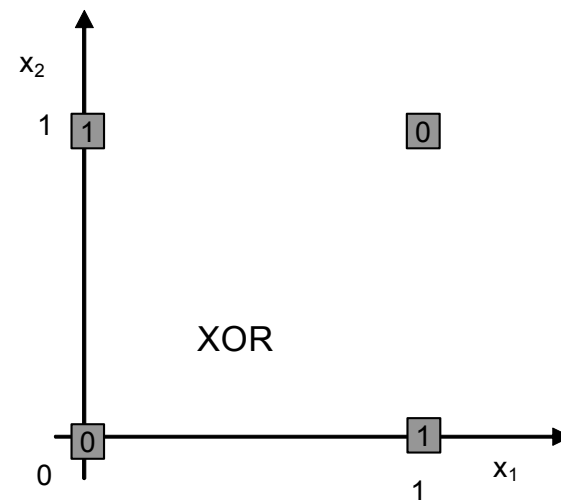
❖ Ένα perceptron με n γραμμές εισόδου μπορεί να θεωρηθεί ότι αναπαριστά ένα υπερεπίπεδο $n-1$ διαστάσεων που διαχωρίζει τα διανύσματα εισόδου σε δύο ομάδες, ανάλογα με την έξοδο.

❖ γραμμικώς διαχωρίσιμα προβλήματα (*linearly separable*).

❖ Παραδείγματα:



Συνάρτηση AND
(γραμμικώς διαχωρίσιμη)

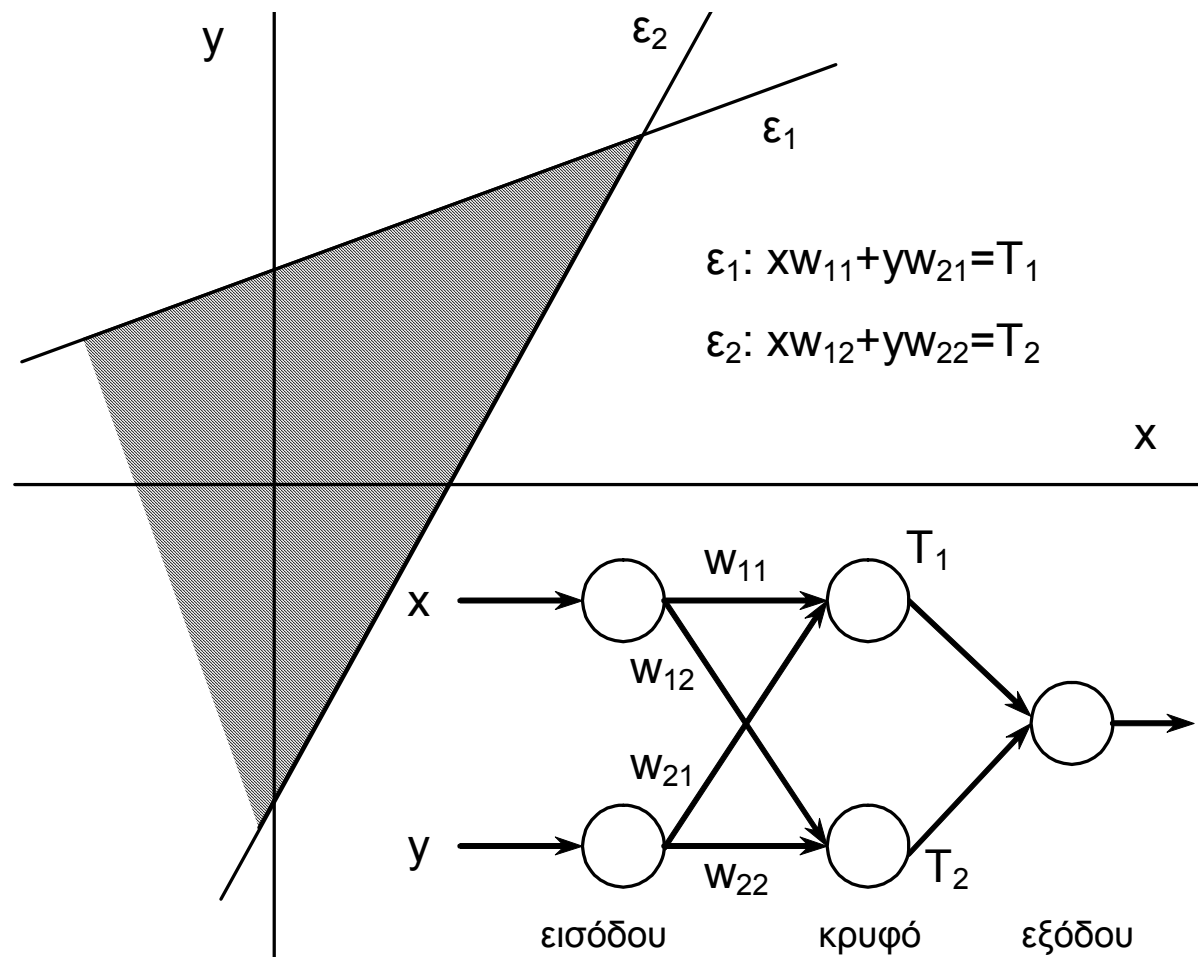


Συνάρτηση XOR
(μη γραμμικώς διαχωρίσιμη)

❖ Τα μη γραμμικώς διαχωρίσιμα προβλήματα απαιτούν τη χρήση ΤΝΔ με ενδιάμεσα κρυφά επίπεδα

Γραμμική Διαχωρισιμότητα (2/2)

- ❖ Γιατί τα ΤΝΔ με κρυφά επίπεδα (και νευρώνες) έχουν καλύτερη διαχωρισιμότητα;
- Ο χώρος των τιμών εισόδου οριοθετείται με πιο πολύπλοκο τρόπο εξαιτίας της πολυπλοκότητας των συνδέσεων.



Κανόνας Δέλτα (1/3)

- ❖ γενίκευση του αλγορίθμου εκπαίδευσης του perceptron
 - ❑ ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος των διανυσμάτων εκπαίδευσης
 - ❑ δεν μπορεί να εφαρμοστεί σε δίκτυα με κρυφά επίπεδα
- ❖ Μέσο τετραγωνικό σφάλμα E στο στοιχειώδες perceptron, για p διανύσματα εκπαίδευσης:

$$E = \frac{1}{p} \sum_{k=1}^p (t_k - input_k)^2$$

- ❑ $input_k$: σήμα εισόδου του νευρώνα (χωρίς συνάρτηση ενεργοποίησης)
 - Συνολικό σήμα εισόδου για κάποιο διάνυσμα εκπαίδευσης p :

$$input = \sum_{i=1}^n w_i x_i$$

- ❑ t_k : επιθυμητή έξοδος

Κανόνας Δέλτα (2/3)

- ❖ Ο κανόνας Δέλτα ακολουθεί την αρνητική κλίση της επιφάνειας σφάλματος, με κατεύθυνση προς το ελάχιστό της:

$$\Delta w_i \propto -\frac{\partial E}{\partial w_i}$$

- ❖ Η παράγωγος του E ως προς τα w είναι:

$$\nabla E = \left(\frac{\partial E}{\partial w_1}, \dots, \frac{\partial E}{\partial w_n} \right)$$

- ❖ Η μεταβολή στην τιμή του βάρους w_i , εξαιτίας της εκπαίδευσης με ένα μόνο από τα διανύσματα εκπαίδευσης, δίνεται από τη σχέση:

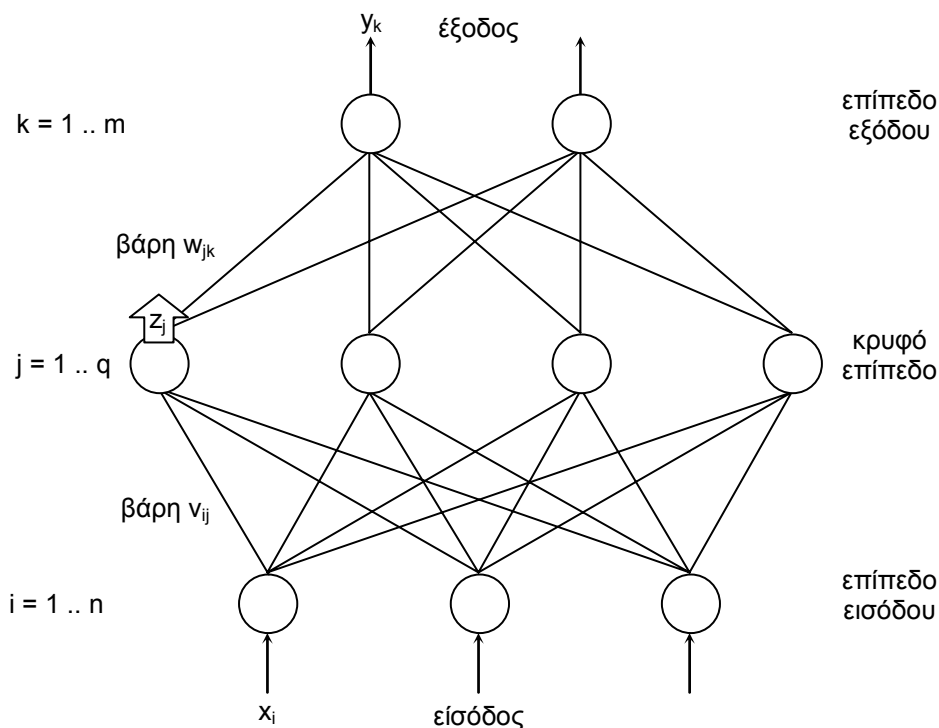
$$\Delta w_i = w_{i(new)} - w_{i(old)} = d(t - input)x_i$$

Ανάστροφη Μετάδοση Λάθους

Back Propagation (1/2)

❖ Βασίζεται στο γενικευμένο κανόνα Δέλτα - ΓΚΔ (*generalized Delta rule*)

❑ **Βασική Ιδέα:** να καθοριστεί το ποσοστό του συνολικού σφάλματος που αντιστοιχεί σε κάθε νευρώνα, ακόμη και αυτών που ανήκουν σε κρυφά επίπεδα.



$$input_j = \sum_{i=1}^n v_{ij} x_i \quad z_j = f(input_j) = f\left(\sum_{i=1}^n v_{ij} x_i\right)$$

$$input_k = \sum_{j=1}^q w_{jk} z_j \quad y_k = f(input_k) = f\left(\sum_{j=1}^q w_{jk} z_j\right)$$

❖ Αποδεικνύεται ότι (σχέσεις ΓΚΔ):

❑ για επίπεδο εξόδου: $\Delta w_{jk} = d \cdot \delta_k \cdot z_j$ με $\delta_k = (t_k - y_k) f'(input_k)$

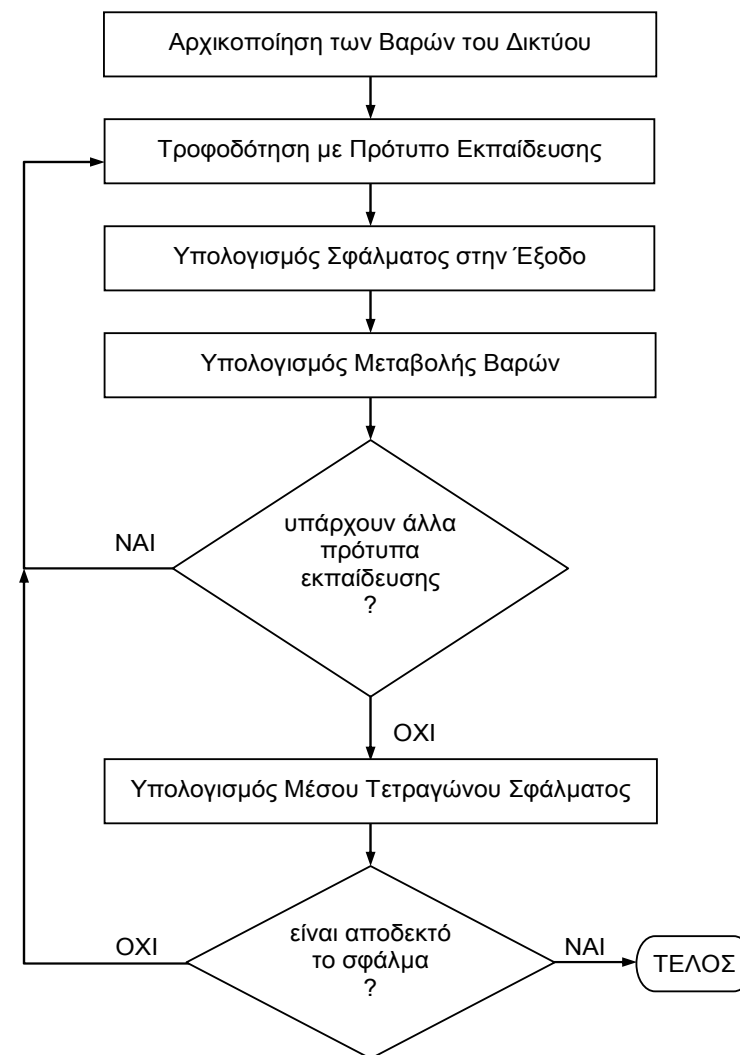
❑ για κρυφό επίπεδο: $\delta_j = f'(input_j) \sum_{k=1}^m \delta_k w_{jk}$

με $\Delta w_{ij} = d \cdot \delta_j \cdot x_i$

Ανάστροφη Μετάδοση Λάθους

Back Propagation (2/2)

- ❖ Η αναπροσαρμογή των βαρών γίνεται από το επίπεδο εξόδου προς το εισόδο.
 - ❑ *ανάστροφο πέρασμα (backward pass) ή ανάστροφη μετάδοση (back propagation)*
- ❖ Διαδικασία βελτιστοποίησης επικλινούς καθόδου (*gradient descent optimization procedure*) που ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα E μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου, για τα p διανύσματα εκπαίδευσης.





Ανάστροφη Μετάδοση Λάθους

Προβλήματα

- ❖ τοπικά ελάχιστα
- ❖ *network paralysis*:
 - ❑ ένα ή περισσότερα βάρη έχουν σταθερά υψηλές απόλυτες τιμές και δεν τροποποιούνται σημαντικά σε κάθε διόρθωση

Παρατηρήσεις

- ❖ τα διανύσματα εκπαίδευσης δεν πρέπει να μεταβάλλονται: δεν επέρχεται σύγκλιση.
- ❖ ο έλεγχος απόδοσης γίνεται με *πρότυπα αξιολόγησης (validation data)*.



Μνήμες Συσχέτισης

Associative Memories

- ❖ Συστήματα μνήμης που ορίζουν απεικονίσεις μεταξύ δύο αναπαραστάσεων X και Y έτσι ώστε όταν δοθεί η μία να μπορεί να ανακληθεί η άλλη.
- ❖ Ανάλογα με τις διαφορές μεταξύ εισόδου και εξόδου διακρίνουμε:
 - αυτοσυσχετιζόμενες μνήμες (*auto-associative memories*)
 - ετεροσυσχετιζόμενες μνήμες (*hetero-associative memories*)
- ❖ Ανάλογα με το αν η έξοδός τους μπορεί να είναι προϊόν παρεμβολής, διακρίνουμε:
 - με δυνατότητα παρεμβολής (*interpolative associative memories*)
 - προσαυξητική μνήμη συσχέτισης (*accretive associative memory*)
- ❖ Τύποι ΤΝΔ που συνιστούν μνήμες συσχέτισης:
 - Γραμμικοί Συσχετιστές, Δίκτυα Hopfield, Μνήμες Συσχέτισης Διπλής Κατεύθυνσης

Γραμμικοί Συσχετιστές (*linear associator*) (1/3)

- ❖ ΤΝΔ πρόσθιας τροφοδότησης, γραμμική συνάρτηση ενεργοποίησης
- ❖ Πρότυπα συσχέτισης: δυαδικά (binary, 0 ή/και 1) ή διπολικά (bipolar, +1 ή/και -1)
- ❖ Εκπαίδευση:

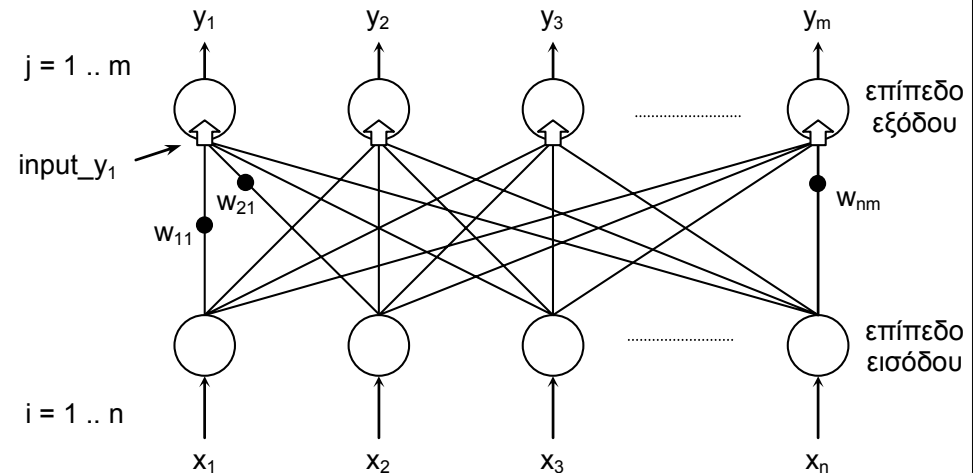
- ❑ $\Delta w_{ij} = x_i \cdot y_j$ (κανόνας Hebb)
- ❑ μετά από p ζευγάρια εκπαίδευσης

$$w_{ij} = \sum_{k=1}^p w_{ki} w_{kj}$$

- ❑ εκφυλίζεται σε πολ/σμό πινάκων

Αν $X = (x_1, x_2, \dots, x_n)$, $Y = (y_1, y_2, \dots, y_m)$

$$W_k = X_k^T \cdot Y_k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot [y_1 \ y_2 \ \dots \ y_m] = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_m \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_m \\ \vdots & \vdots & \ddots & \vdots \\ x_n y_1 & x_n y_2 & \dots & x_n y_m \end{bmatrix} = \begin{bmatrix} \Delta w_{11} & \Delta w_{12} & \dots & \Delta w_{1m} \\ \Delta w_{21} & \Delta w_{22} & \dots & \Delta w_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta w_{n1} & \Delta w_{n2} & \dots & \Delta w_{nm} \end{bmatrix}$$



- ❖ Για όλο το σύνολο εκπαίδευσης: $W = \sum_{k=1}^p W_k = \sum_{k=1}^p X_k^T \cdot Y_k$

- ❖ Για κάθε X:Y το W αναλύεται σε: $f_1(X, Y) + f_2(\text{υπόλοιπων X:Y})$ ($2^{\text{ος}}$ όρος: **crosstalk**)

Γραμμικοί Συσχετιστές (2/3)

- ❖ Ανάκληση: είσοδος ένα διάνυσμα X - αναμενόμενη έξοδος το συσχετιζόμενο διάν. Y
- ❖ Στην πράξη, ο υπολογισμός μπορεί να γίνει με χρήση πινάκων: $Y_k = f(X_k \cdot W)$
 - όπου f είναι η συνάρτηση ενεργοποίησης των νευρώνων του επιπέδου εξόδου

$$y_j = \begin{cases} 1 & , y_{input_j} > 0 \\ 0 & , y_{input_j} \leq 0 \end{cases}$$

← για δυαδικά διανύσματα X και Y

για διπολικά διανύσματα X και $Y \Rightarrow$

$$y_j = \begin{cases} +1 & , y_{input_j} \geq 0 \\ -1 & , y_{input_j} < 0 \end{cases}$$

- y_{input_j} είναι το συνολικό σήμα εισόδου στο νευρώνα j του επιπέδου εξόδου

- ❖ Παράδειγμα ετεροσυσχετιζόμενης μνήμης (με ΤΝΔ 4-2):

Δεδομένα Εκπαίδευσης

	X_i	Y_j
p_1 :	(1, 0, 0, 0)	(1, 0)
p_2 :	(1, 1, 0, 0)	(1, 0)
p_3 :	(0, 0, 0, 1)	(0, 1)
p_4 :	(0, 0, 1, 1)	(0, 1)

$$W_1 = X_1^T \cdot Y_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot [1 \ 0] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$W_2 = X_2^T \cdot Y_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot [1 \ 0] = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$W_3 = X_3^T \cdot Y_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \cdot [0 \ 1] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$W_4 = X_4^T \cdot Y_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \cdot [0 \ 1] = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$W = \sum_{k=1}^4 W_k = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}$$

- ❖ Ανάκληση: έστω $X_i = p_1$. Θα ανακληθεί το: $Y_1 = f \left([1 \ 0 \ 0 \ 0] \cdot \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 2 \end{bmatrix} \right) = f([2 \ 0]) = [1 \ 0]$ (σωστό)

Γραμμικοί Συσχετιστές (3/3)

- ❖ Παράδειγμα αυτοσυσχετιζόμενης μνήμης: ζητούμενο είναι ένας γραμμικός συσχετιστής που να ανακαλεί το διπολικό διάνυσμα $X=(1, 1, 1, -1)$. (προφανώς με ΤΝΔ 4-4)

- ❖ Πίνακας Βαρών: $W = X^T \cdot X = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \cdot [1 \ 1 \ 1 \ -1] = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix}$

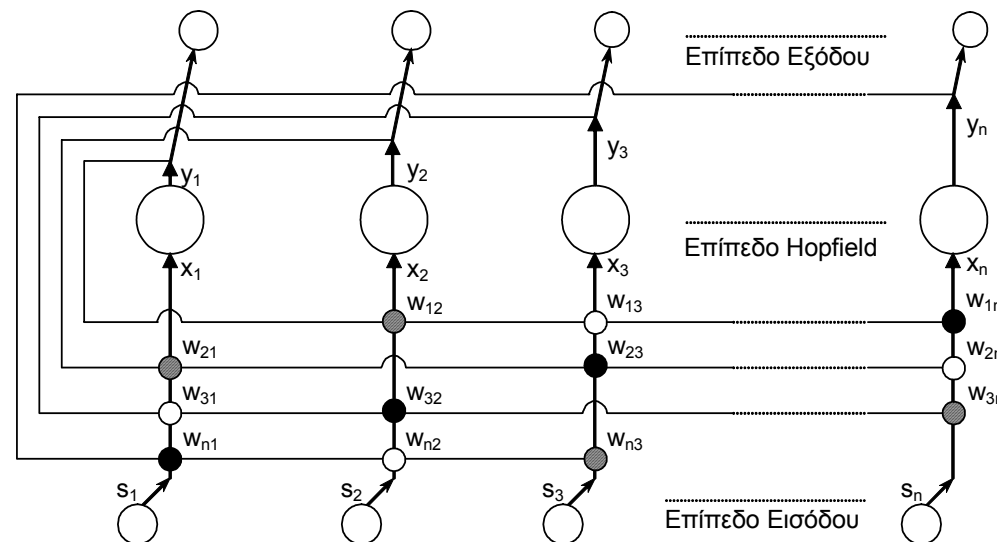
- ❖ Ανάκληση: ανοχή σε θόρυβο και σε ελλιπή δεδομένα

αρχικό διάνυσμα εκπαίδευσης:	$[1 \ 1 \ 1 \ -1] \cdot W = [4 \ 4 \ 4 \ -4] \rightarrow [1 \ 1 \ 1 \ -1]$
αρχικό διάνυσμα με θόρυβο:	$[-1 \ 1 \ 1 \ -1] \cdot W = [2 \ 2 \ 2 \ -2] \rightarrow [1 \ 1 \ 1 \ -1]$
αρχικό ελλιπές διάνυσμα:	$[0 \ 0 \ 1 \ -1] \cdot W = [2 \ 2 \ 2 \ -2] \rightarrow [1 \ 1 \ 1 \ -1]$
αρχικό διάνυσμα με πολύ θόρυβο:	$[-1 \ -1 \ 1 \ -1] \cdot W = [0 \ 0 \ 0 \ 0] \rightarrow \text{δεν γίνεται αναγνώριση}$

- ❖ Καθώς αποθηκεύονται όλο και περισσότερα πρότυπα, το δίκτυο χάνει τη **δυνατότητα διόρθωσης**: γιατί ο πίνακας W γίνεται ταυτοτικός πίνακας (*identity matrix*).
- ❖ Η **τέλεια ανάκληση** είναι δυνατή μόνο όταν τα διανύσματα εισόδου είναι ορθογώνια μεταξύ τους, δηλαδή όταν $X_i \cdot X_j = 0$ για $i \neq j$.
- ❖ Χωρητικότητα (με τέλεια ανάκληση) δομής $n \times n$: $n-1$ διπολικά ζευγάρια

Δίκτυα Hopfield (1/3)

- ❖ Ένα μόνο επίπεδο νευρώνων (Hopfield)
- ❖ Δίκτυα με ανατροφοδότηση (recurrent): αμφίδρομες συνδέσεις, συμμετρικά βάρη
- ❖ αυτοσυσχετιζόμενες μνήμες
- ❖ υπολογίζουν την έξοδό τους αναδρομικά στο χρόνο, μέχρις ότου το σύστημα να σταθεροποιηθεί



Διανύσματα εισόδου: $S=(s_1, s_2, \dots, s_n)$	Είσοδοι σε επίπεδο Hopfield	Για p πρότυπα προς αποθήκευση:	Συναρτ. Ενεργοποίησης
Έξοδος του δικτύου: (y_1, y_2, \dots, y_n)	$x_j = \sum_{i=1, i \neq j}^n y_i w_{ij} + s_j$	$W = \sum_{k=1}^p W_k = \sum_{k=1}^p S_k^T \cdot S_k$	$y_i = \begin{cases} 1 & , x_i > \theta_i \\ y_i' & , x_i = \theta_i \\ -1 & , x_i < \theta_i \end{cases}$

- ❖ Ο πίνακας βαρών W είναι τετράγωνος, συμμετρικός ($w_{ij}=w_{ji}$), με $w_{ii}=0$
- ❖ Στη συνάρτηση ενεργοποίησης, y_i' είναι η έξοδος του προηγούμενου κύκλου υπολογισμών, ενώ συνήθως $\theta_i=0$.

Δίκτυα Hopfield (2/3)

❖ **Παράδειγμα:** θέλουμε να αποθηκευτεί το διάνυσμα (πρότυπο) $S_I=(1, 1, 1, 0)$

❖ Πίνακας Βαρών: $W_1 = S_1^T \cdot S_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ -1 \end{bmatrix} \cdot [1 \ 1 \ 1 \ -1] = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$

- ❑ έγινε μετατροπή του δυαδικού διανύσματος S_I σε διπολικό (αντικατάσταση των 0 με -1)
- ❑ τέθηκε $w_{ii}=0$, όπως ορίζει η δομή του δικτύου Hopfield

❖ **Ανάκληση:** για κάθε διάνυσμα εισόδου, το δίκτυο θα ισορροπήσει στο "κοντινότερο" διάνυσμα από το σύνολο των διανυσμάτων με τα οποία εκπαιδεύτηκε.

❖ Διαδικασία ανάκλησης (τερματίζει όταν το δίκτυο συγκλίνει, κάτι εξασφαλισμένο):

1. Δίνεται στο δίκτυο το δυαδικό διάνυσμα εισόδου S'
2. Αρχικοποιείται το δίκτυο με $x_i=S_i$, $y_i=x_i$ και $convergence=false$
3. Εφόσον το δίκτυο δεν συγκλίνει, δηλ. εφόσον $convergence=false$
 - α) Επιλέγεται τυχαία ένας από τους νευρώνες, έστω ο j

$$x_j = \sum_{i=1, i \neq j}^n y_i w_{ij} + s_j$$

β) Υπολογίζεται η είσοδος του

γ) Εφαρμόζεται η συνάρτηση ενεργοποίησης του νευρώνα j στην είσοδο x_j

δ) Περιοδικά, ελέγχεται εάν έχει επέλθει σύγκλιση (υπολογισμός $convergence$)

Δίκτυα Hopfield (3/3)

- ❖ Έστω το δίκτυο Hopfield με τον πίνακα βαρών W_1 (προηγούμενη διαφάνεια)
- ❖ Τροφοδοτείται με το διάνυσμα $S'=(s'_1, s'_2, s'_3, s'_4)=(0, 0, 1, 0)$, αλλοιωμένη μορφή του S_1 που είναι αποθηκευμένο στο δίκτυο.
- ❖ Αρχικά η έξοδος είναι $Y=S'=(0, 0, 1, 0)$ και $\theta_i=0$ ($i=1,2,3,4$)
- ❖ Κύκλοι λειτουργίας:
 - Έστω ότι επιλέγεται ο νευρώνας $j=1$:
 - $x_1 = \sum y_i \cdot w_{i1} + s'_1 = (0 \cdot 0 + 0 \cdot 1 + 1 \cdot 1 + 0 \cdot (-1)) + 0 = 1$
Επειδή $1 > \theta_1$ θα είναι $y_1=1$. Άρα η έξοδος γίνεται $Y=(1, 0, 1, 0)$
 - Έστω τώρα ότι επιλέγεται ο νευρώνας $j=4$.
 - $x_4 = \sum y_i \cdot w_{i4} + s'_4 = (1 \cdot (-1) + 0 \cdot (-1) + 1 \cdot (-1) + 0 \cdot 0) + 0 = -2$
Επειδή $-2 < \theta_4$ θα είναι $y_4=-1$. Άρα η έξοδος είναι $(1, 0, 1, -1)$ και σε δυαδική μορφή $(1, 0, 1, 0)$.
 - Επιλέγεται αυτή τη φορά ο νευρώνας $j=3$.
 - $x_3 = \sum y_i \cdot w_{i3} + s'_3 = (1 \cdot 1 + 0 \cdot 1 + 1 \cdot 0 + 0 \cdot (-1)) + 1 = 2$
Επειδή $2 > \theta_3$ θα είναι $y_3=1$. Άρα η έξοδος παραμένει $(1, 0, 1, 0)$
 - Τέλος, επιλέγεται ο νευρώνας $j=2$
 - $x_2 = \sum y_i \cdot w_{i2} + s'_2 = (1 \cdot 1 + 0 \cdot 0 + 1 \cdot 1 + 0 \cdot (-1)) + 0 = 2$
Επειδή $2 > \theta_2$ θα είναι $y_2=1$. Άρα η έξοδος γίνεται $(1, 1, 1, 0)$ **σωστή ανάκληση**

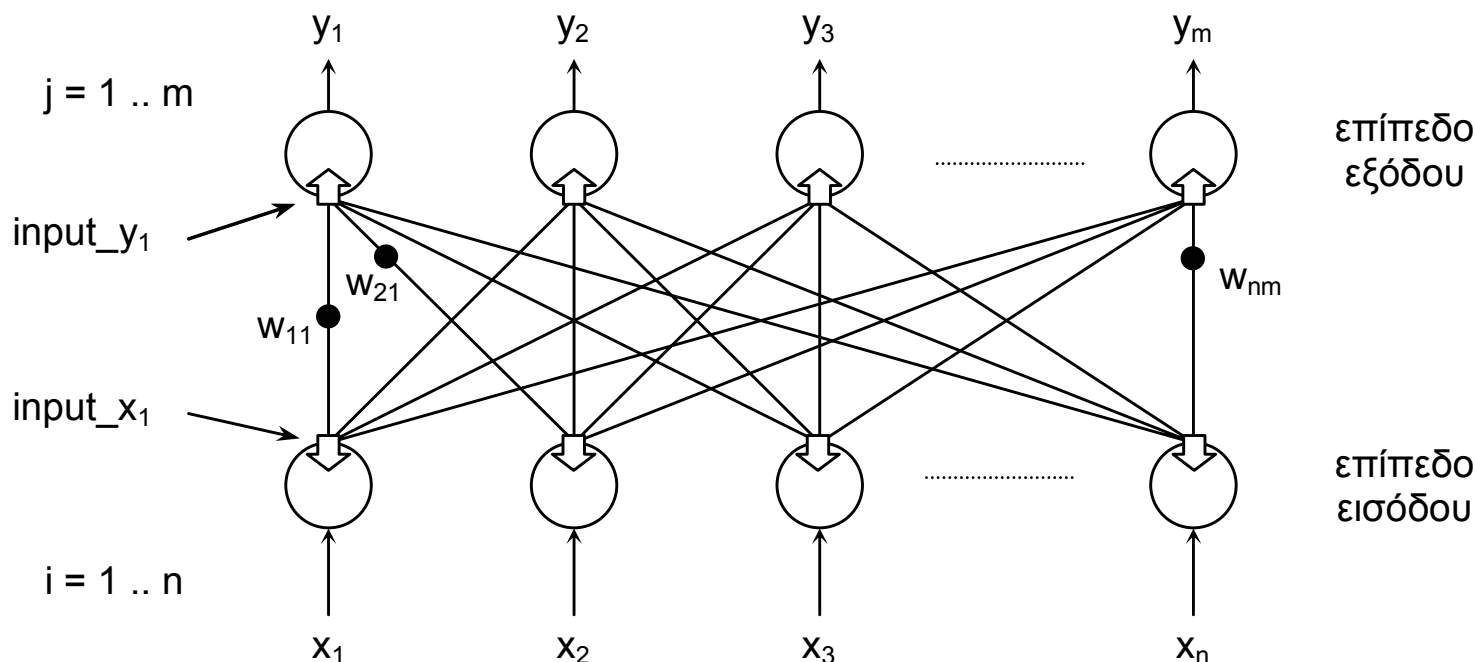
$$y_i = \begin{cases} 1 & , x_i > \theta_i \\ y_i' & , x_i = \theta_i \\ -1 & , x_i < \theta_i \end{cases}$$

Σημείωση: στα Σ χρησιμοποιείται το τρέχων Y σε δυαδική μορφή (0 και 1 αντί -1 και 1) και το διπολικό αποτέλεσμα που παράγεται μετατρέπεται σε δυαδικό για να χρησιμοποιηθεί στο επόμενο βήμα.

Μνήμες Συσχέτισης Διπλής Κατεύθυνσης

(Bidirectional Associative Memories - BAM)

- ❖ Επέκταση μοντέλου Hopfield με ένα επιπλέον επίπεδο νευρώνων.
- ❖ Οι συνδέσεις μεταξύ των νευρώνων είναι διπλής κατεύθυνσης, δηλαδή $w_{ij}=w_{ji}$



- ❖ Οι νευρώνες στα δύο επίπεδα λειτουργούν σαν στοιχεία εισόδου ή εξόδου, ανάλογα με την κατεύθυνση προς την οποία γίνονται υπολογισμοί.
 - ❑ Συνάρτηση Ενεργοποίησης: διακριτή συνάρτηση κατωφλίου.
 - ❑ Διανύσματα X και Y : διπολικά ή δυαδικά.

Μνήμες Συσχέτισης Διπλής Κατεύθυνσης

- ❖ Εκπαίδευση: με ζευγάρια διπολικών διανυσμάτων εισόδου-εξόδου και βηματική συνάρτηση ενεργοποίησης.

πίνακας βαρών	W για ταυτόχρονη αποθήκευση p ζευγαριών	είσοδος σε κάθε νευρώνα του επιπέδου Y	έξοδος y_j
$W_k = X_k^T \cdot Y_k$	$W = \sum_{k=1}^p W_k$	$input_y_j = \sum_{i=1}^n x_i w_{ij}$	$y_j = \begin{cases} 1 & , input_y_j > 0 \\ y_j' & , input_y_j = 0 \\ -1 & , input_y_j < 0 \end{cases}$

- ❖ Ανάκληση: έστω X^* το διάνυσμα εισόδου (πιθανώς με θόρυβο)

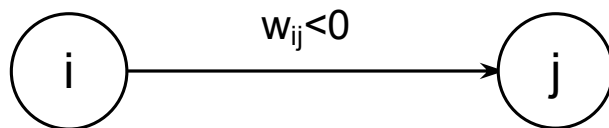
1. Δίνεται στο δίκτυο το διάνυσμα εισόδου X^* . Επίσης τίθεται *convergence=false*.
2. Υπολογίζεται το διάνυσμα εξόδου Y' του δικτύου από τη σχέση $Y' = M^T \times X^*$.
3. Εφόσον το δίκτυο δεν συγκλίνει, δηλ. εφόσον *convergence=false*
 - α) Γίνεται ανάδραση και υπολογίζεται το $X' = M^T \times Y'$.
 - β) Υπολογίζεται το νέο διάνυσμα εξόδου Y'' από τη σχέση $Y'' = M^T \times X'$.
 - γ) Γίνεται έλεγχος αν έχει επέλθει σύγκλιση

□ σύγκλιση (και ανάκληση) όταν δύο διαδοχικές έξοδοι στο βήμα 3β είναι ίδιες

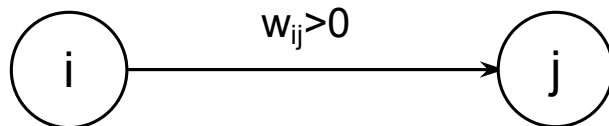
- ❖ χωρητικότητα μοντέλου BAM: $O(\max(n, m))$.

ΤΝΔ με Ανταγωνισμό

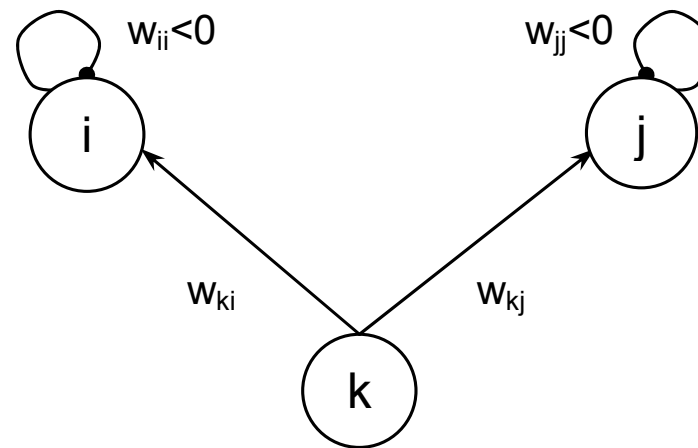
- ❖ **Βασική Ιδέα:** οι νευρώνες πρέπει να είναι σε θέση να επηρεάσουν θετικά, ουδέτερα ή ακόμη και αρνητικά τους υπόλοιπους νευρώνες του δικτύου.
 - ❑ ποιος νευρώνας θα ανταποκριθεί περισσότερο;
 - ❑ απλούστερη περίπτωση: μόνο ο νευρώνας με τη μεγαλύτερη έξοδο (νικητής) παράγει τελικά αποτέλεσμα (*winner-takes-all* - WTA).
- ❖ Μοντελοποίηση Ανταγωνισμού
 - ❑ **Παράπλευρη καταστολή** ή **ενίσχυση** (*lateral inhibition* ή *excitation*)
 - ❑ **Ανταγωνισμός πόρων** (*resource competition*)
 - βιολογικά αποδεκτό - μοντελοποιεί και το βιολογικό φαινόμενο της εξασθένησης (*self decay*).



Παράπλευρη Καταστολή



Παράπλευρη Ενίσχυση



Ανταγωνισμός Πόρων

Δίκτυα Kohonen (1/5)

- ❖ *Αυτο-οργανούμενη Απεικόνιση Χαρακτηριστικών - Self-organizing Feature Map*
- ❖ Οι νευρώνες είναι τοποθετημένοι σε κάποια γεωμετρική τοπολογία (πχ επίπεδο, κτλ).
- ❖ Κάθε νευρώνας είναι συνδεδεμένος μέσω βαρών με την είσοδο που αποτελείται από k στοιχεία και λαμβάνει ένα πλήρες αντίγραφο του διανύσματος εισόδου τροποποιημένου από τα βάρη.
- ❖ Μέσω εκπαίδευσης, μαθαίνουν να αντιστοιχούν ένα σήμα εισόδου με συγκεκριμένο νευρώνα στο επίπεδο εξόδου: άρα πραγματοποιούν κατηγοριοποίηση.
- ❖ Μάθηση χωρίς επίβλεψη.
- ❖ Μερικοί τρόποι υλοποίησης ανταγωνισμού για επιλογή του "νικητή νευρώνα":
 - ❑ Μέσω συνάρτησης μεγίστου προσδιορίζεται ο νευρώνας με τη μεγαλύτερη απόκριση στην είσοδο και στη συνέχεια αποδίδεται έξοδος +1 σε αυτόν και μηδέν στους υπόλοιπους.
 - ❑ Ο νευρώνας με τη μεγαλύτερη τιμή στο εσωτερικό γινόμενο μεταξύ του διανύσματος εισόδου και του διανύσματος των βαρών.
 - ❑ Ο νευρώνας για τον οποίο η Ευκλείδεια απόσταση μεταξύ του διανύσματος εισόδου και του διανύσματος των βαρών είναι η μικρότερη.

Δίκτυα Kohonen (2/5)

❖ Εκπαίδευση δικτύου Kohonen με k στοιχεία εισόδου.

- ❑ ανάθεση τυχαίων μικρών τιμών στα βάρη εισόδου
- ❑ συνήθως, τα αρχικά βάρη και τα διανύσματα εκπαίδευσης κανονικοποιούνται ως προς 1

- σχέση κανονικοποίησης για διάνυσμα εισόδου $S \implies$
- όμοια κανονικοποιούνται και τα αρχικά βάρη εισόδου $W_i = (w_{i1}, w_{i2}, \dots, w_{ik})$

$$s'_i = \frac{s_i}{[s_1^2 + s_2^2 + \dots + s_k^2]^{1/2}}$$

- ❑ για κάθε νευρώνα του επιπέδου Kohonen, υπολογίζεται το πόσο κοντά στο διάνυσμα εκπαίδευσης βρίσκονται τα βάρη εισόδου του νευρώνα

$$d_i = \sqrt{\sum_{k=1}^n (s_k - w_{ik})^2}$$

- ❑ ο νευρώνας c με το μικρότερο d είναι ο "νικητής" και θα "εκπαιδευτεί"

$$\Delta w_{ij} = \begin{cases} n(s_j - w_{ij}) & \text{αν ο νευρώνας } i \text{ ανήκει στη γειτονιά του } c \\ 0 & \text{αν ο νευρώνας } i \text{ δεν ανήκει στη γειτονιά του } c \end{cases}$$

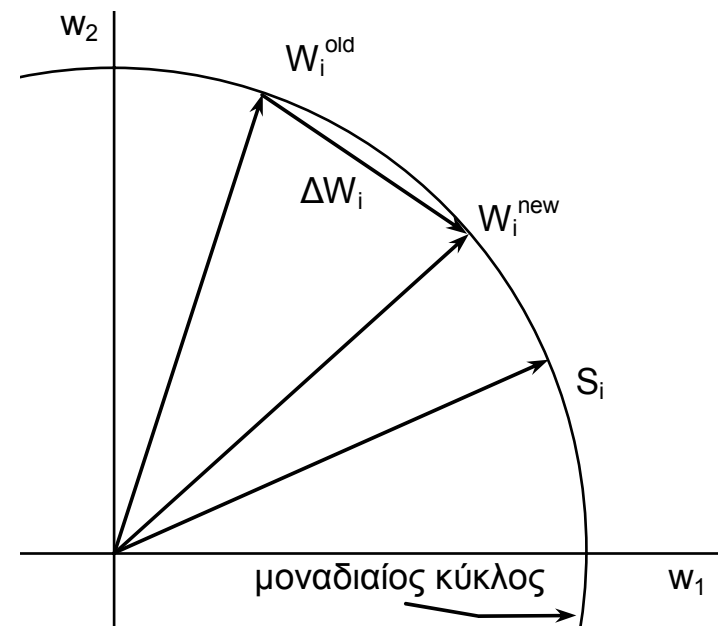
❖ επόμενος κύκλος: ο ρυθμός εκπαίδευσης n και το μέγεθος της γειτονιάς ελαττώνονται

Δίκτυα Kohonen (3/5)

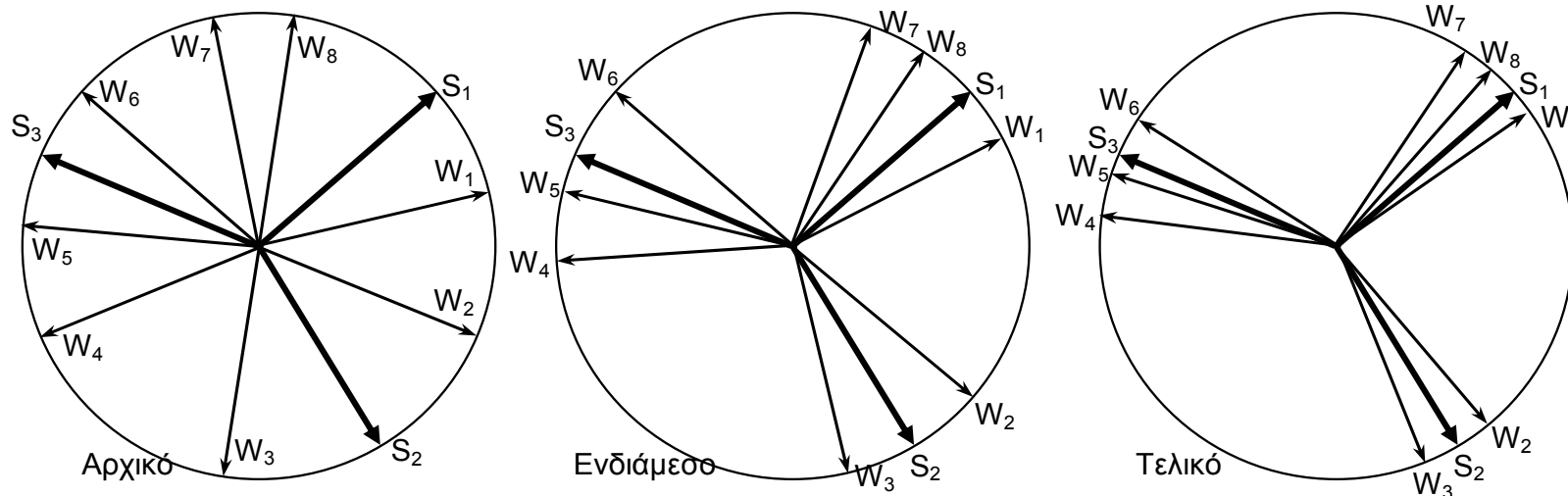
❖ Βήμα εκπαίδευσης σε δίκτυο Kohonen με 2 στοιχεία εισόδου:

❖ Ο αριθμός των νευρώνων εξόδου επηρεάζει:

- ❑ μικρός: μπορεί να οδηγήσει σε συγχώνευση κλάσεων
- ❑ μεγάλος: οδηγεί σε υπερπροσαρμογή (περισσότερες κατηγορίες από όσες υπάρχουν)



❖ Εξέλιξη εκπαίδευσης σε νευρωνικό δίκτυο Kohonen



Δίκτυα Kohonen (4/5)

❖ Κακή αρχικοποίηση τιμών για τα βάρη (σχήμα).

❖ Παράδειγμα εκπαίδευσης και ανάκλησης:

❑ Έστω δίκτυο Kohonen 4-2 και τα 4 διανύσματα εκπαίδευσης: $S_1=(1,1,0,0)$, $S_2=(0,0,0,1)$, $S_3=(1,0,0,0)$ και $S_4=(0,0,1,1)$

❑ αρχικά διανύσματα βαρών: $W_1=(0.2, 0.6, 0.5, 0.9)$ και $W_2=(0.8, 0.4, 0.7, 0.3)$

❑ αρχικός ρυθμός εκπαίδευσης: $n_0=0.6$

❑ μεταβολή n με το χρόνο εκπαίδευσης (εποχές): $n_{(t+1)}=0.5n_{(t)}$

❑ δίνουμε στην είσοδο το S_1 και υπολογίζουμε την Ευκλείδεια απόσταση μεταξύ S_1 και W_1, W_2

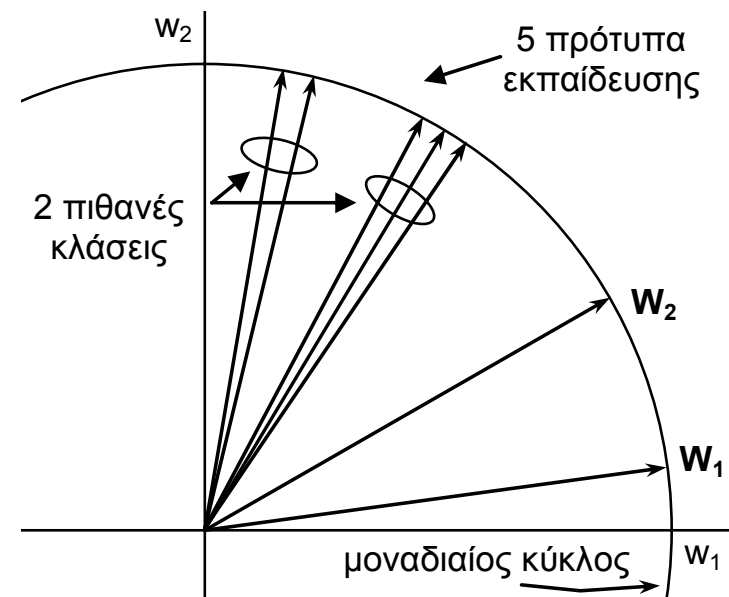
- $d_1 = \sqrt{(s_1 - w_1)^2 + (s_2 - w_2)^2 + (s_3 - w_3)^2 + (s_4 - w_4)^2} = \sqrt{(1 - 0.2)^2 + (1 - 0.6)^2 + (0 - 0.5)^2 + (0 - 0.9)^2} = 1.3638$

- όμοια: $d_2=0.9899$

❑ Επειδή $d_2 < d_1$, μόνο το διάνυσμα βαρών W_2 θα μεταβληθεί λόγω εκπαίδευσης

- Για τη συνιστώσα w_1 του W_2 : $w_{1new} = w_{1old} + n(s_1 - w_{1old}) = 0.8 + 0.6(1 - 0.8) = 0.92$ (όμοια τα υπόλοιπα)

- Τελικά: $W_{2new} = (0.92, 0.76, 0.28, 0.12)$ ενώ $W_{1new} = W_{1old}$.



Δίκτυα Kohonen (5/5)

- ❖ Η διαδικασία επαναλαμβάνεται για τα υπόλοιπα διανύσματα εκπαίδευσης.
 - ❑ μετά το πέρας ενός κύκλου εκπαίδευσης:
 $W_1=(0.032, 0.096, 0.680, 0.984)$ και $W_2=(0.968, 0.304, 0.112, 0.048)$
- ❖ Στη συνέχεια, μεταβάλλεται ο ρυθμός εκπαίδευσης σε 0.3 και αρχίζει νέα εποχή.
 - ❑ μετά το πέρας του δεύτερου κύκλου εκπαίδευσης:
 $W_1=(0.016, 0.047, 0.633, 0.992)$ και $W_2=(0.984, 0.359, 0.055, 0.024)$
 - ❑ μετά 10 κύκλους εκπαίδευσης:
 $W_1=(0.0, 0.0, 0.5, 1.0)$ και $W_2=(1.0, 0.5, 0.0, 0.0)$
- ❖ **Ανάκληση**
 - ❑ Ο προσδιορισμός της κλάσης ($1^{\text{η}}$ ή $2^{\text{η}}$) στην οποία ανήκει κάθε διάνυσμα S_i γίνεται μέσω του εσωτερικού γινομένου $W_i \cdot S_j$.
 - ❑ Το μεγαλύτερο εσωτερικό γινόμενο $W_i \cdot S_j$ για δεδομένο j καθορίζει τη νικητήρια κλάση, που προσδιορίζεται από το i .
 - $j=1$ $W_1 \cdot S_1=0$ και $W_2 \cdot S_1=1.5$ άρα το S_1 ανήκει στην $2^{\text{η}}$ κλάση
 - $j=2$ $W_1 \cdot S_2=1$ και $W_2 \cdot S_2=0$ άρα το S_2 ανήκει στην $1^{\text{η}}$ κλάση
 - $j=3$ $W_1 \cdot S_3=0$ και $W_2 \cdot S_3=1$ άρα το S_3 ανήκει στην $2^{\text{η}}$ κλάση
 - $j=4$ $W_1 \cdot S_4=1.5$ και $W_2 \cdot S_4=0$ άρα το S_4 ανήκει στην $1^{\text{η}}$ κλάση



Εφαρμογές Νευρωνικών Δικτύων

- ❖ Δημοφιλή σε προβλήματα που περιέχουν μη-προβλέψιμες λειτουργίες και τα οποία δεν είναι πλήρως κατανοητά.

- ❖ Κατηγοριοποίηση
 - ❑ Ιατρική, Άμυνα, Γεωργία, Οικονομία/επιχειρήσεις

- ❖ Αναγνώριση
 - ❑ Τράπεζες, Πληροφορική και Τηλεπικοινωνίες

- ❖ Αποτίμηση
 - ❑ Άμυνα, Ασφάλεια, Μηχανολογία

- ❖ Πρόβλεψη
 - ❑ Οικονομία/επιχειρήσεις, Γεωργία, Μετεωρολογία