



Αριστοτέλειο  
Πανεπιστήμιο  
Θεσσαλονίκης

# Τεχνητή Νοημοσύνη

Τεχνητά Νευρωνικά Δίκτυα

Ιώαννης Βλαχάβας

Τμήμα Πληροφορικής ΑΠΘ



## Άδειες Χρήσης

Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons. Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



## Χρηματοδότηση

Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα. Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.



Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.





# *Τεχνητά Νευρωνικά Δίκτυα* *(Artificial Neural Nets)*

# Εισαγωγή

- ❖ Τα νευρωνικά δίκτυα (ΝΔ) είναι μια ιδιαίτερη προσέγγιση στη δημιουργία συστημάτων με νοημοσύνη καθώς αποφεύγουν να αναπαραστήσουν ρητά τη γνώση και να υιοθετήσουν ειδικά σχεδιασμένους αλγόριθμους αναζήτησης.
- ❖ Βασίζονται σε βιολογικά πρότυπα καθώς χρησιμοποιούν δομές και διαδικασίες που μιμούνται τις αντίστοιχες του ανθρώπινου εγκεφάλου.

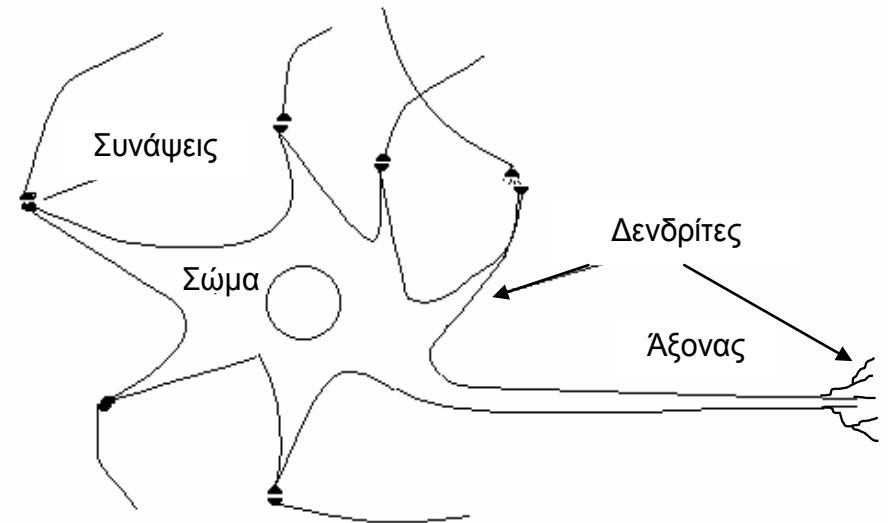
## Βιολογικός Νευρώνας

- ❖ *Νευρώνας*: Η δομική μονάδα του εγκεφάλου.

- ❑ *Σώμα*: Ο πυρήνας του νευρώνα.
- ❑ *Δενδρίτες*: Μεταφέρουν ηλεκτρικά σήματα από γειτονικούς νευρώνες (σημεία εισόδου).
- ❑ *Άξονας*: Υπό κατάλληλες συνθήκες, εξάγει ηλεκτρικά σήματα προς γειτονικούς νευρώνες.

- ❖ Στο άκρο κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό που ονομάζονται *σύναψη*.

- ❑ Η ικανότητα μάθησης και μνήμης που παρουσιάζει ο εγκέφαλος οφείλεται στην ικανότητα των συνάψεων να μεταβάλουν την αγωγιμότητά τους.
- ❑ Τα σήματα που εισέρχονται στο *σώμα* μέσω των *δενδριτών*, συνδυάζονται και αν το αποτέλεσμα ξεπερνά κάποιο *κατώφλι*, διαδίδεται μέσω του *άξονα* προς άλλους νευρώνες.

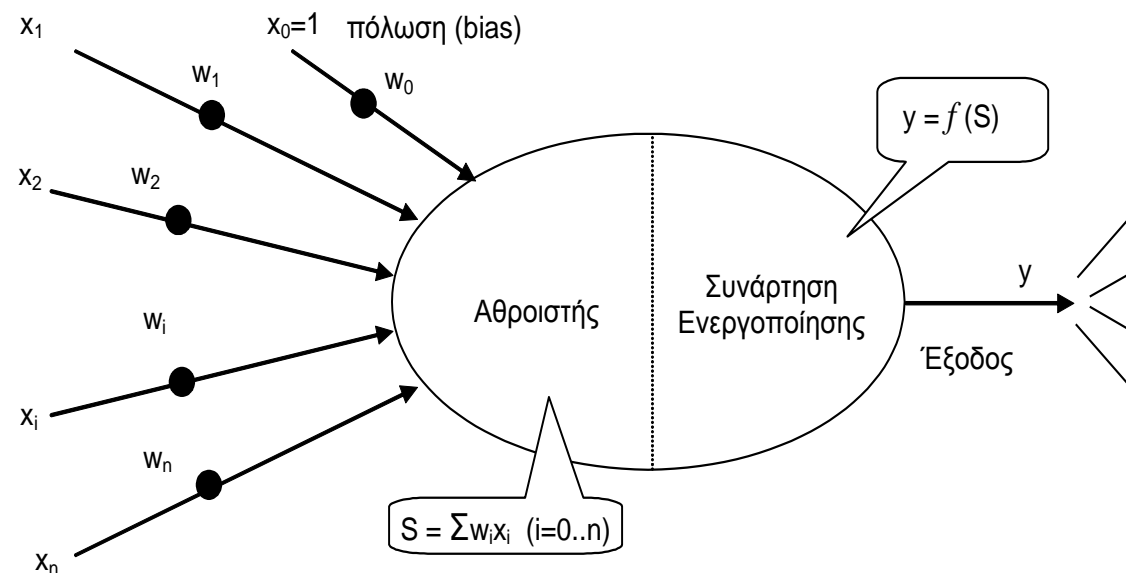


# Φυσικά Νευρωνικά Δίκτυα

- ❖ Ο εγκέφαλος του ανθρώπου αποτελείται από περίπου 100 δισεκατομμύρια νευρώνες.
  - ❑ Κάθε νευρώνας συνδέεται κατά μέσο όρο με 1.000 άλλους νευρώνες.
  - ❑ Άρα, υπάρχουν συνολικά περίπου 100 τρισεκατομμύρια συνάψεις που επηρεάζουν την λειτουργία του εγκεφάλου.
- ❖ Είναι προφανές ότι κάθε προσπάθειά να αντιγραφεί η δομή και η λειτουργία του εγκεφάλου σε τέτοια κλίμακα είναι αδύνατη.
  - ❑ Τα μοντέλα που κατασκευάζονται περιλαμβάνουν μερικές χιλιάδες τεχνητούς νευρώνες, έχουν το πολύ ένα εκατομμύριο τεχνητές συνάψεις και παρουσιάζουν πολύ περιορισμένη λειτουργικότητα σε σχέση με τον ανθρώπινο εγκέφαλο.
- ❖ Αν και ο χρόνος απόκρισης των βιολογικών νευρώνων είναι της τάξης των χιλιοστών του δευτερολέπτου (msec), εντούτοις ο εγκέφαλος είναι σε θέση να λαμβάνει πολύπλοκες αποφάσεις, εκπληκτικά γρήγορα.
- ❖ Κατά μία άποψη, αυτό οφείλεται στο ότι η υπολογιστική ικανότητα του εγκεφάλου και η πληροφορία που περιέχει είναι διαμοιρασμένα σε όλο του τον όγκο.
  - ❑ Πρόκειται δηλαδή για ένα παράλληλο και κατανεμημένο υπολογιστικό σύστημα.
- ❖ Αυτά τα χαρακτηριστικά είναι και το κυριότερο κίνητρο πίσω από την επιθυμία να μοντελοποιηθεί ο ανθρώπινος εγκέφαλος με τα **τεχνητά νευρωνικά δίκτυα**.

# Μοντέλο Τεχνητού Νευρώνα (*artificial neuron*)

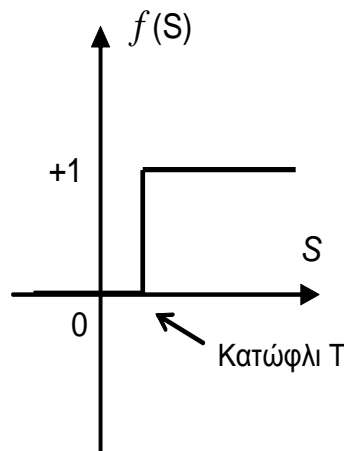
- ❖ Υπολογιστικό μοντέλο με δομή αντίστοιχη του βιολογικού νευρώνα.
- ❖ Δέχεται *σήματα εισόδου*  $x_1, x_2, \dots, x_n$ : Συνεχείς μεταβλητές (αντί ηλεκτρικών παλμών).
- ❖ Κάθε σήμα εισόδου μεταβάλλεται από μια αρνητική ή θετική *τιμή βάρους*  $w_i$  (*weight*): Αντίστοιχο των συνάψεων.
  - ❑ *Πόλωση* (*bias*): Ειδική περίπτωση βάρους ( $w_0$ ) που επιδρά σε τιμή εισόδου  $x_0=1$ . Πρόκειται για εξωτερικό ερέθισμα. Μπορεί να παρέχει έμμεσο έλεγχο στη συνάρτηση ενεργοποίησης.
- ❖ Το σώμα του τεχνητού νευρώνα αποτελείται από:
  - ❑ τον *αθροιστή* (*sum*): προσθέτει τα επηρεασμένα από τα βάρη σήματα εισόδου και παράγει την ποσότητα  $S$ .
  - ❑ τη *συνάρτηση ενεργοποίησης ή κατωφλίου* (*activation ή threshold function*): μη γραμμικό φίλτρο που διαμορφώνει το σήμα εξόδου  $y$ , σε συνάρτηση με την ποσότητα  $S$ .
- ❖ *Έξοδος* (αριθμητική τιμή)
  - ❑ (Μπορεί να) Αποτελεί είσοδο σε άλλους νευρώνες.



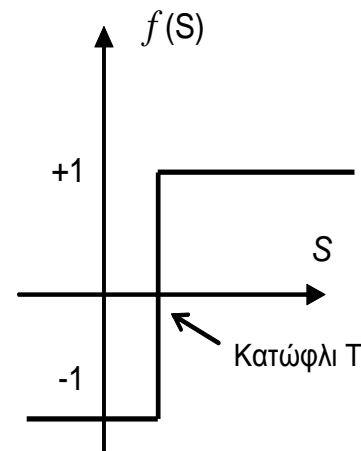
# Συναρτήσεις Ενεργοποίησης

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

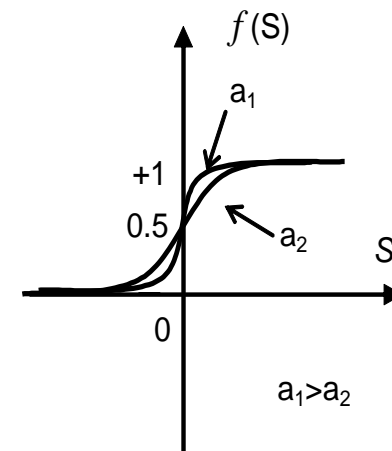
❖ Η βηματική (*step*) συνάρτηση: δίνει στην έξοδο αποτέλεσμα (συνήθως 1) μόνο αν η τιμή που υπολογίζει ο αθροιστής είναι μεγαλύτερη από μία τιμή κατωφλίου  $T$ .



α) Βηματική Συνάρτηση



β) Συνάρτηση Προσήμου



γ) Λογιστική Συνάρτηση

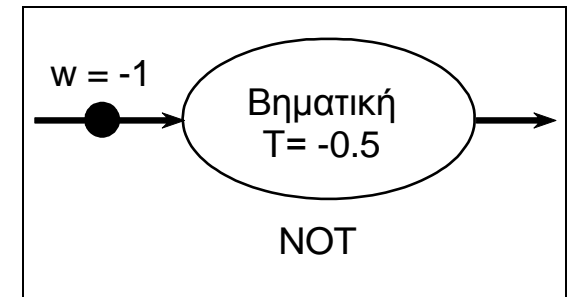
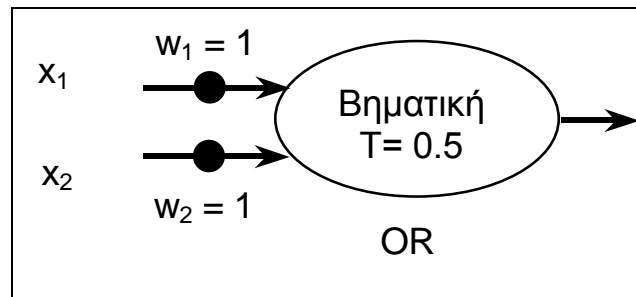
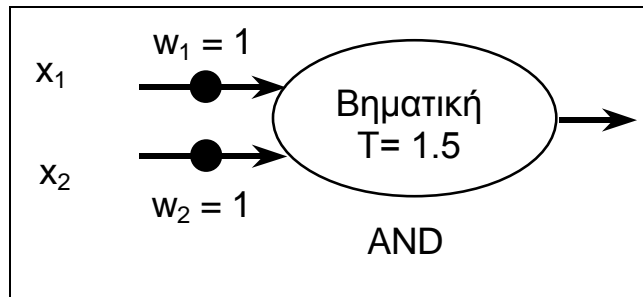
❖ Η συνάρτηση πρόσημου (*sign*) δίνει αρνητική (ή θετική) έξοδο αν η τιμή που υπολογίζει ο αθροιστής είναι μικρότερη (ή μεγαλύτερη) από μία τιμή κατωφλίου  $T$ .

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

- ❑ γραφική παράσταση σχήματος πεπλατυσμένου "S"
- ❑ ο συντελεστής  $a$  ρυθμίζει την ταχύτητα μετάβασης μεταξύ των δύο ακραίων τιμών (συνήθως 0 και 1 ή -1 και 1)
- ❑ συνεχείς και παραγωγίσιμες σε όλο το φάσμα τιμών εισόδου
- ❑ λειτουργούν ως φίλτρο που καταστέλλει τις μεγάλες τιμές

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

# Υλοποίηση Λογικών Συναρτήσεων με Τεχνητό Νευρώνα



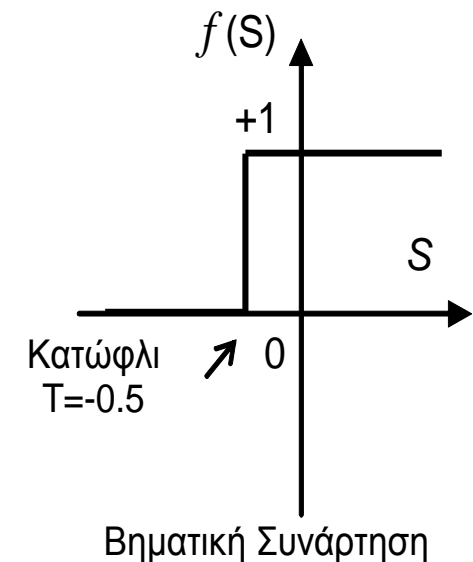
## ❖ Παράδειγμα: υλοποίηση του NOT:

### □ Σχεδίαση:

- Συνάρτηση Ενεργοποίησης: η βηματική συνάρτηση με κατώφλι  $T=-0.5$
- Οι τιμές εισόδου μπορούν να είναι 0 (ψευδές) ή 1 (αληθές).

### □ Λειτουργία:

- Αν η είσοδος του νευρώνα είναι 0 τότε πολλαπλασιάζεται με το βάρος  $w=-1$  δίνει  $S=0$ . Η τιμή αυτή ξεπερνά το κατώφλι του  $-0.5$ , οπότε στην έξοδο παράγεται 1.
- Στην περίπτωση που η τιμή εισόδου είναι 1 τότε  $S = -1$ , τιμή που βρίσκεται κάτω του κατωφλίου  $-0.5$ , με αποτέλεσμα να παράγεται στην έξοδο 0.





# Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)

- ❖ Συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος τεχνητών νευρώνων οργανωμένων σε δομές παρόμοιες με αυτές του ανθρώπινου εγκεφάλου.
- ❖ Αποτελούνται από στρώματα ή επίπεδα (layers) νευρώνων

- ❑ *επίπεδο εισόδου (input layer)*

- ✓ περιέχει νευρώνες χωρίς συνάρτηση ενεργοποίησης και βάρη εισόδου
- ✓ δεν εκτελούν κάποιο υπολογισμό
- ✓ προωθούν κατάλληλα το σήμα εισόδου σε επόμενο επίπεδο

- ❑ *ενδιάμεσα ή κρυφά επίπεδα (hidden layers)*

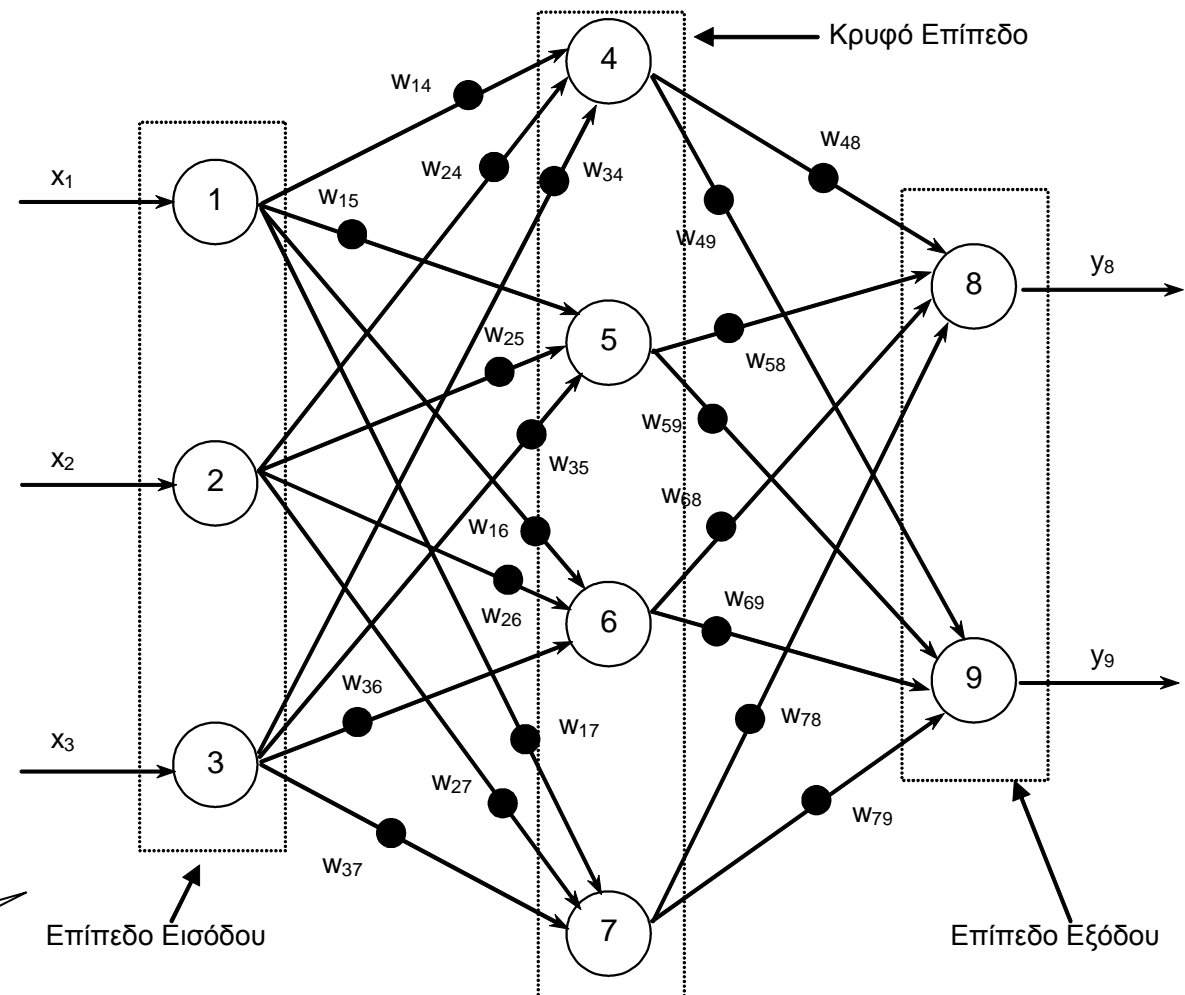
- ✓ Προαιρετικά. Μπορούν να υπάρχουν, ένα ή περισσότερα.

- ❑ *επίπεδο εξόδου (output layer)*

- ❖ Συντομογραφία για πολυεπίπεδα

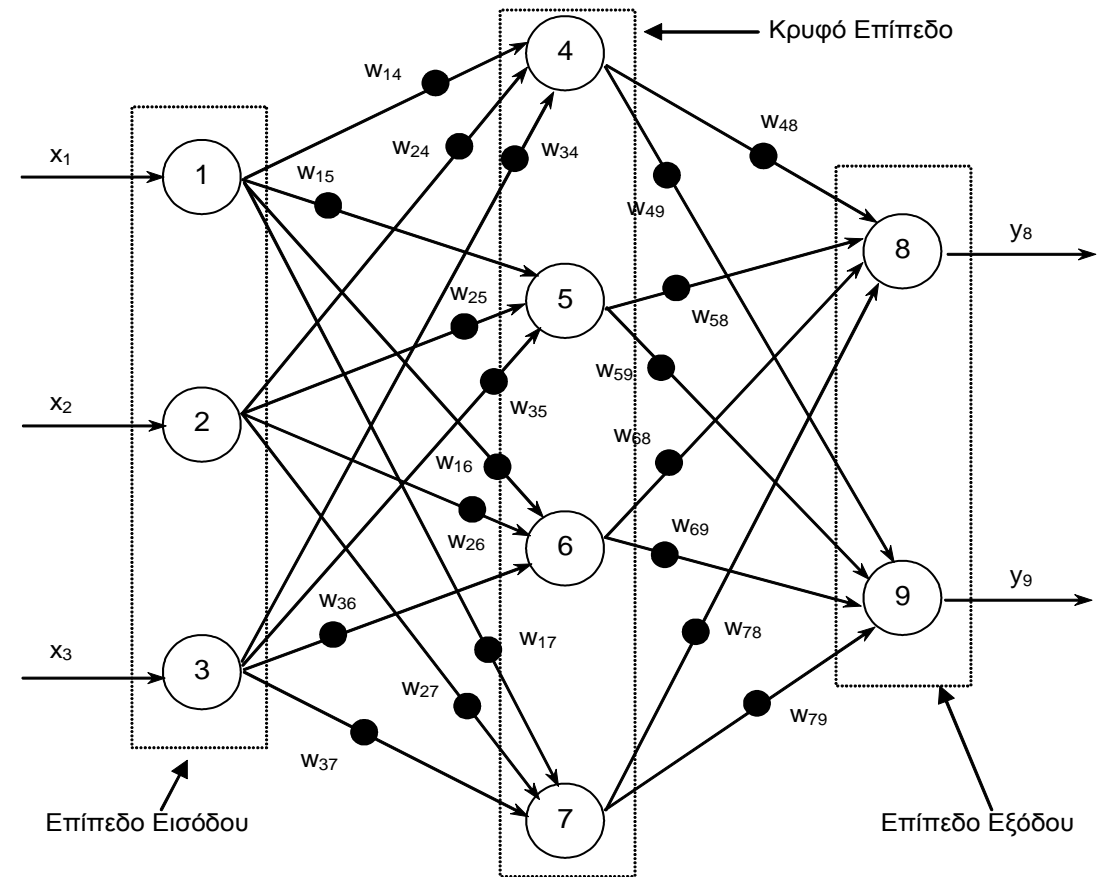
ΤΝΔ:  $(p, m_1, m_2, \dots, m_q, n)$

Πλήρως συνδεδεμένο ΤΝΔ απλής τροφοδότησης 3-4-2.



# Ροή Πληροφορίας

- ❖ Το ΤΝΔ του σχήματος είναι ένα πλήρως συνδεδεμένο πρόσθιας τροφοδότησης (feedforward).
- ❑ Δέχεται τα σήματα εισόδου  $x_1$ ,  $x_2$  και  $x_3$  (διάνυσμα εισόδου) και μετά από εσωτερική επεξεργασία παράγει στην έξοδο τα σήματα  $y_8$  και  $y_9$  (διάνυσμα εξόδου).
- ❑ Για παράδειγμα, ο νευρώνας 5 δέχεται τρία σήματα εισόδου από τους νευρώνες (1, 2 και 3) του επιπέδου εισόδου.
- ❑ Τα σήματα αυτά τροποποιούνται από τα βάρη  $w_{15}$ ,  $w_{25}$  και  $w_{35}$  και αθροίζονται από τον αθροιστή του νευρώνα 5.
- ❑ Το συνολικό σήμα από τον αθροιστή δέχεται τη δράση της συνάρτησης ενεργοποίησης του νευρώνα 5 και το αποτέλεσμα στέλνεται στους νευρώνες 8 και 9.
- ❑ Η λειτουργία των υπολοίπων νευρώνων είναι παρόμοια.



# Χαρακτηριστικά - Ορολογία

- ❖ Οι νευρώνες των διαφόρων στρωμάτων μπορεί να είναι:
  - ❑ Πλήρως συνδεδεμένοι (*fully connected*)
    - Συνδέονται με όλους τους νευρώνες του επόμενου επιπέδου.
  - ❑ Μερικώς συνδεδεμένοι (*partially connected*)
- ❖ Τα ΤΝΔ χαρακτηρίζονται ως:
  - ❑ Δίκτυα με πρόσθια τροφοδότηση (*feedforward*)
    - Δεν υπάρχουν συνδέσεις από νευρώνες ενός επιπέδου προς νευρώνες προηγούμενου επιπέδου.
    - Η ροή πληροφορίας είναι μιας κατεύθυνσης.
  - ❑ Δίκτυα με ανατροφοδότηση (*feedback ή recurrent*)
    - Υπάρχουν συνδέσεις από νευρώνες ενός επιπέδου προς νευρώνες προηγούμενου επιπέδου ή/και συνδέσεις μεταξύ νευρώνων ίδιου επιπέδου.
    - Οι υπολογισμοί γίνονται σε δύο στάδια:
      - ✓ στο πρώτο στάδιο υπολογίζονται τα μεγέθη που αφορούν συνδέσεις πρόσθιας τροφοδότησης
      - ✓ στο δεύτερο γίνονται οι υπολογισμοί για τις συνδέσεις ανατροφοδότησης.
- ❖ Αν και σε ορισμένες περιπτώσεις τα δίκτυα με ανατροφοδότηση είναι πολύ χρήσιμα, στην πλειοψηφία των εφαρμογών νευρωνικών δικτύων χρησιμοποιούνται δίκτυα πρόσθιας τροφοδότησης.

# Μάθηση (Εκπαίδευση) και Ανάκληση

- ❖ Τα ΤΝΔ πραγματοποιούν δύο βασικές λειτουργίες: τη μάθηση και την 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 ένα συγκεκριμένο διάνυσμα εισόδου και συγκεκριμένες τιμές βαρών.
  
- ❖ Ο γενικός τρόπος με τον οποίο γίνεται η τροποποίηση των βαρών ενός ΤΝΔ κατά την εκπαίδευσή του, επιτρέπει τη διάκριση τριών ειδών μάθησης:
  - ❑ **Μάθηση υπό Επίβλεψη** (*supervised learning*)
  - ❑ **Μάθηση χωρίς Επίβλεψη** (*unsupervised learning*)
  - ❑ **Βαθμολογημένη Μάθηση** (*graded learning*)

# Είδη Μάθησης

## ❖ *Μάθηση υπό επίβλεψη (supervised learning):*

- ❑ Στο δίκτυο δίνονται ζευγάρια διανυσμάτων εισόδου - επιθυμητής εξόδου.
- ❑ Το ΤΝΔ, με την τρέχουσα κατάσταση βαρών, παράγει μία έξοδο η οποία αρχικά διαφέρει από την επιθυμητή έξοδο.
- ❑ Αυτή η διαφορά ονομάζεται *σφάλμα (error)* και βάσει αυτής καθώς και ενός αλγορίθμου εκπαίδευσης γίνεται συνήθως η αναπροσαρμογή των βαρών.

## ❖ *Μάθηση χωρίς επίβλεψη (unsupervised learning)*

- ❑ Η απόκριση του δικτύου βασίζεται στην ικανότητά του να αυτο-οργανώνεται με βάση τα διανύσματα εισόδου (π.χ. δίκτυα Kohonen).
- ❑ Αυτή η εσωτερική οργάνωση γίνεται έτσι ώστε σε συγκεκριμένο σύνολο εισόδων να αντιδρά ισχυρά ένας συγκεκριμένος νευρώνας.
- ❑ Τέτοια σύνολα εισόδων, αντιστοιχούν σε έννοιες και χαρακτηριστικά του πραγματικού κόσμου τα οποία το ΤΝΔ καλείται να μάθει.

## ❖ *Βαθμολογημένη μάθηση (graded learning)*

- ❑ Η έξοδος χαρακτηρίζεται ως "καλή" ή "κακή" με βάση μια αριθμητική κλίμακα και τα βάρη αναπροσαρμόζονται με βάση αυτό το χαρακτηρισμό.

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*)

- ❑ Η διαφορά μεταξύ πραγματικής και επιθυμητής εξόδου ελαχιστοποιείται μέσω μιας διαδικασίας ελαχίστων τετραγώνων.
  - Δε μπορεί να εφαρμοστεί σε δίκτυα με κρυφά επίπεδα επειδή δεν είναι γνωστή η επιθυμητή έξοδος για τους νευρώνες των κρυφών επιπέδων.

## ❖ Αλγόριθμος ανάστροφης μετάδοσης λάθους (*back propagation*)

- ❑ Η μεταβολή των βαρών βασίζεται στον υπολογισμό της συνεισφοράς κάθε βάρους στο συνολικό σφάλμα.

## ❖ Ανταγωνιστική μάθηση (*competitive learning*)

- ❑ Οι τεχνητοί νευρώνες συναγωνίζονται, κατά κάποιο τρόπο, μεταξύ τους και μόνο αυτός με τη μεγαλύτερη απόκριση σε δοθείσα είσοδο τροποποιεί τα βάρη του.

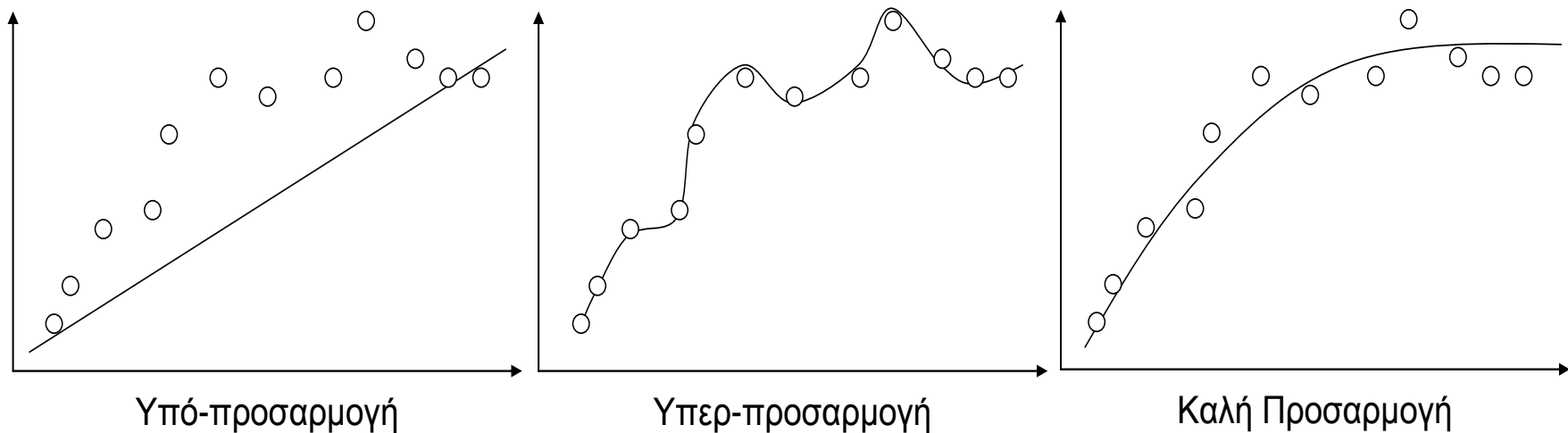
## ❖ Τυχαία μάθηση (*random learning*)

- ❑ Οι μεταβολές στα βάρη εισάγονται τυχαία και ανάλογα με το αν η έξοδος βελτιώνεται ή όχι με βάση κάποια προκαθορισμένα από το χρήστη κριτήρια, οι μεταβολές αυτές υιοθετούνται ή απορρίπτονται.

# Χαρακτηριστικά Εκπαίδευσης

❖ Στην εκπαίδευση, μπορεί να εμφανιστούν φαινόμενα:

- ❑ υποπροσαρμογής ή ατελούς μάθησης (*underfitting*): αποτυχία μοντελοποίησης δεδομένων εκπαίδευσης, π.χ. λόγω απλότητας του ΤΝΔ
- ❑ υπερπροσαρμογής (*overfitting*): ένα πολύπλοκο ΤΝΔ ενδέχεται να μοντελοποιήσει υπερβολικά τα δεδομένα εκπαίδευσης καθώς και το θόρυβο που πιθανώς υπάρχει σε αυτά (απομνημόνευση)
  - θα δίνει σωστή πρόβλεψη για τα δεδομένα εκπαίδευσης αλλά λανθασμένες προβλέψεις για άλλα δεδομένα εισόδου.



❖ Σε ΤΝΔ με κρυφά επίπεδα (εμπειρικές εκτιμήσεις):

- ❑ δεδομένα εκπαίδευσης με θόρυβο: τουλάχιστο 30 φορές περισσότερα από τα βάρη
- ❑ δεδομένα εκπαίδευσης χωρίς θόρυβο: τουλάχιστο 5 φορές περισσότερα από τα βάρη

# Δεδομένα Εκπαίδευσης

- ❖ Συνηθέστερος τρόπος χρήσης των δεδομένων εκπαίδευσης: σε κύκλους εκπαίδευσης που ονομάζονται *εποχές* (*epochs*). Ειδικότερα:
  - ❑ *μάθηση δέσμης* (*batch learning*)
    - το δίκτυο δέχεται ως είσοδο, ένα-ένα, όλα τα διανύσματα εκπαίδευσης,
    - αθροίζει τη μεταβολή στην τιμή των βαρών που προκύπτει από κάθε διάνυσμα και
    - αναπροσαρμόζει τα βάρη στο τέλος κάθε κύκλου, βάσει της συσσωρευμένης μεταβολής
    - δίνει γρηγορότερα αποτέλεσμα, αλλά έχει μεγαλύτερες απαιτήσεις μνήμης
  - ❑ *επαυξητική μάθηση* (*incremental learning*)
    - όπως η μάθηση δέσμης, αλλά η αναπροσαρμογή των βαρών γίνεται μετά τη χρήση ενός από τα διανύσματα εκπαίδευσης
  - ❑ συνδυασμός των δύο παραπάνω μεθόδων
- ❖ Η εκπαίδευση συνήθως 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).

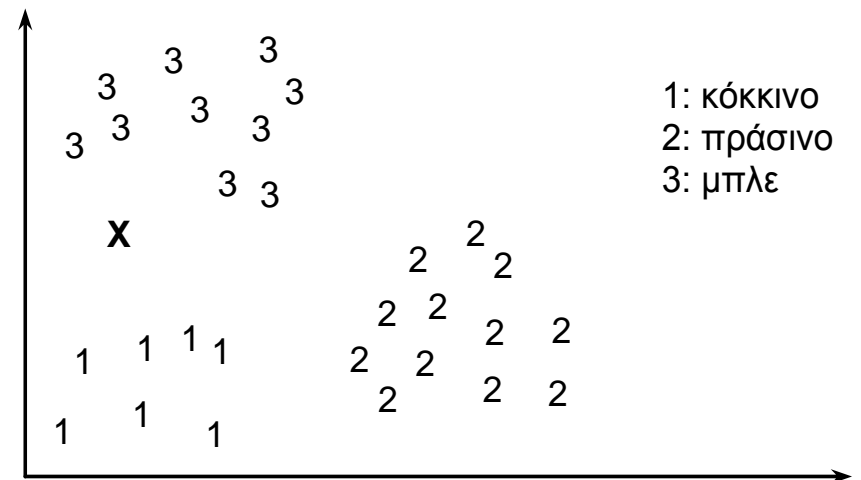
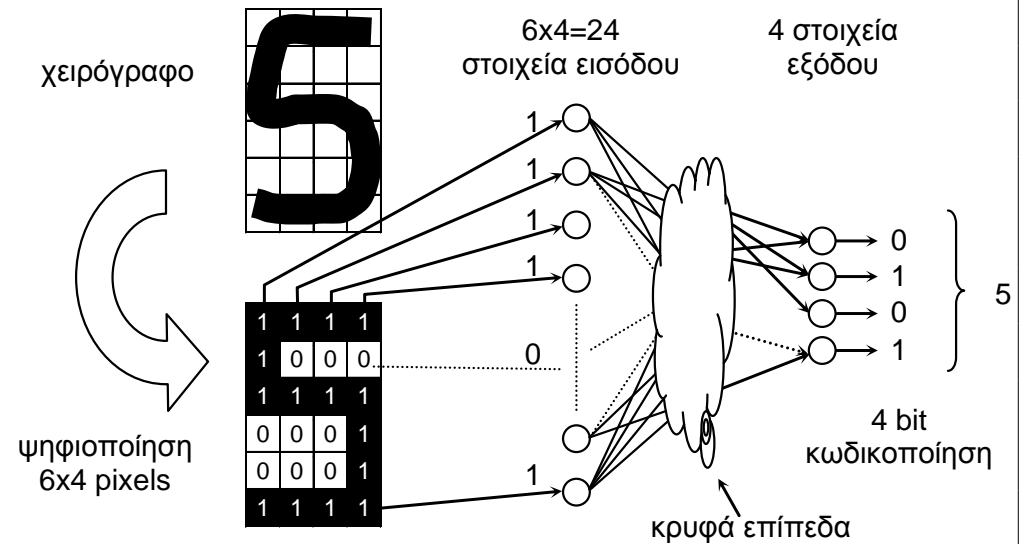


# Βασικές Ιδιότητες των ΤΝΔ

- ❖ **Ικανότητα να μαθαίνουν μέσω παραδειγμάτων (learn by example).**
  - ❑ Οργανώνουν την πληροφορία των δεδομένων εισόδου σε χρήσιμες μορφές.
  - ❑ Κωδικοποιούν τη σχέση που ισχύει μεταξύ των δεδομένων εισόδου και εξόδου.
- ❖ **Η δυνατότητα θεώρησής τους ως κατανεμημένη μνήμη (distributed memory) και ως μνήμη συσχέτισης (associative memory).**
  - ❑ Η πληροφορία που κωδικοποιούν είναι κατανεμημένη σε όλα τα βάρη της συνδεσμολογίας τους.
  - ❑ Η ανάκληση της πληροφορίας γίνεται με βάση το περιεχόμενο και όχι τη διεύθυνση, όπως δηλαδή συμβαίνει και με τον ανθρώπινο εγκέφαλο.
- ❖ **Η μεγάλη τους ανοχή σε σφάλματα (fault-tolerant).**
  - ❑ Η κακή λειτουργία ή η καταστροφή ενός νευρώνα ή κάποιων συνδέσεων δεν είναι ικανή να διαταράξει σημαντικά την λειτουργία τους καθώς η πληροφορία που εσωκλείουν δεν είναι εντοπισμένη σε συγκεκριμένο σημείο αλλά διάχυτη σε όλο το δίκτυο.
  - ❑ Γενικά, το μέγεθος του σφάλματος λόγω "δομικών αστοχιών" είναι ανάλογο του ποσοστού των κατεστραμμένων συνδέσεων.
- ❖ **Η εξαιρετική ικανότητά τους για αναγνώριση προτύπων (pattern recognition).**
  - ❑ Δεν επηρεάζονται από ελλιπή ή/και με θόρυβο δεδομένα.

# ΤΝΔ Πρόσθιας Τροφοδότησης (*feedforward*)

- ❖ Είναι η πιο απλή μορφή νευρωνικών δικτύων.
- ❖ Η ροή πληροφορίας μέσα στο δίκτυο είναι μονής κατεύθυνσης (πρόσθια).
- ❖ Αποτελούνται από ένα επίπεδο εισόδου, ένα επίπεδο εξόδου, κανένα, ένα ή περισσότερα ενδιάμεσα, κρυφά επίπεδα.
- ❖ **Είδος μάθησης:** μάθηση με επίβλεψη.
- ❖ **Τοπολογία** του δικτύου: κρυφά επίπεδα, νευρώνες ανά επίπεδο, συνδεσμολογία.
  - ❑ Γενικά, δεν υπάρχει κανόνας για τον προσδιορισμό αυτών των μεγεθών.
  - ❑ Τα δεδομένα εισόδου-εξόδου βοηθούν στην εκτίμηση του αριθμού νευρώνων στα επίπεδα εισόδου και εξόδου. π.χ.:
    - κλασική κωδικοποίηση με bit
    - ένας νευρώνας εξόδου ανά κατηγορία που θέλουμε να αναγνωρίζει, αρκεί να μην έχουμε κατηγορίες με σειρά (ordered categories)



# 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
  - ❑ λόγος: για να αποφευχθεί η αποκλειστική συσχέτιση ενός κρυφού νευρώνα με ένα διάνυσμα εκπαίδευσης (απομνημόνευση)
- ❖ Συνήθως κάθε νευρώνας συνδέεται με όλους τους νευρώνες του επόμενου επιπέδου.
- ❖ Γενικά, απαιτούνται αρκετές δοκιμές και πειραματισμοί μέχρις ότου βρεθεί μία καλή δομή δικτύου για δεδομένο πρόβλημα.

# Perceptron

- ❖ Η πιο απλή τοπολογία δικτύου με απλή τροφοδότηση.
  - ❑ Πρόκειται για ένα μοναδικό τεχνητό νευρώνα,
  - ❑ Χρησιμοποιεί ως συνάρτηση κατωφλίου τη βηματική συνάρτηση.
- ❖ Μάθηση με επίβλεψη. Ο αλγόριθμος μεταβολής των βαρών είναι ο ακόλουθος:

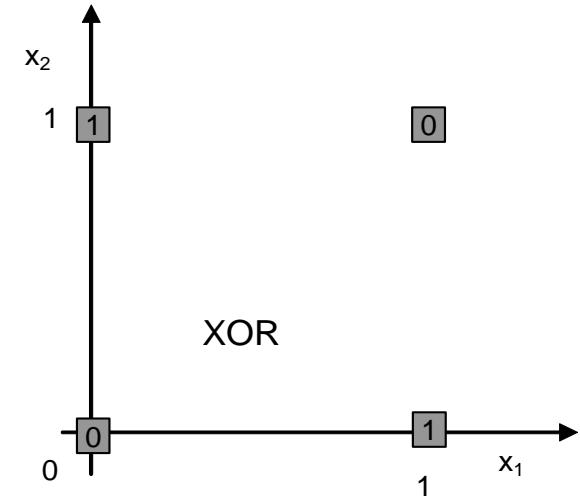
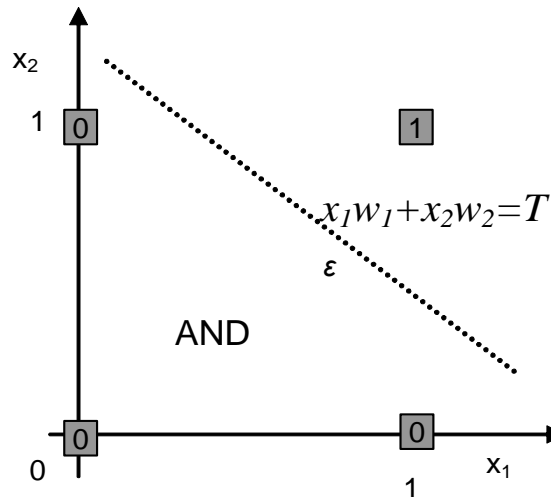
Μέχρις ότου ικανοποιηθεί η συνθήκη τερματισμού της εκπαίδευσης επανέλαβε:

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  $t$ .
- ❖ Η ποσότητα  $d$  ονομάζεται *ρυθμός μάθησης (learning rate)*
  - ❑ έχει συνήθως τιμή μεταξύ 0 και 1
  - ❑ καθορίζει το ρυθμό μεταβολής των βαρών.
- ❖ Αποδεικνύεται ότι εάν υπάρχει ένα διάνυσμα βαρών  $W_1$  που παράγει την επιθυμητή έξοδο για όλα τα διανύσματα εκπαίδευσης, τότε ξεκινώντας από ένα τυχαίο διάνυσμα βαρών  $W_0$  και μετά από πεπερασμένο αριθμό βημάτων, ο αλγόριθμος perceptron θα συγκλίνει σε κάποιο διάνυσμα βαρών  $W^*$ , όχι απαραίτητα το  $W_1$ , το οποίο επίσης θα παράγει την επιθυμητή έξοδο για όλα τα διανύσματα εκπαίδευσης.

# 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, τοποθετώντας από τη μία μεριά όσα παράγουν έξοδο 1 και από την άλλη όσα παράγουν έξοδο 0.
- ❖ Προβλήματα των οποίων οι τιμές εισόδου-εξόδου υπόκεινται σε αυτό τον κανόνα ονομάζονται **γραμμικώς διαχωρίσιμα** (*linearly separable*).
- ❖ Παραδείγματα:



## Συνάρτηση AND

(γραμμικώς διαχωρίσιμη)

Η ευθεία  $\varepsilon$  χωρίζει τα ζευγάρια εισόδου-εξόδου σε δύο περιοχές.

Εκτός από το AND, το perceptron είναι σε θέση να μοντελοποιήσει και τις στοιχειώδεις λογικές πύλες OR και NOT

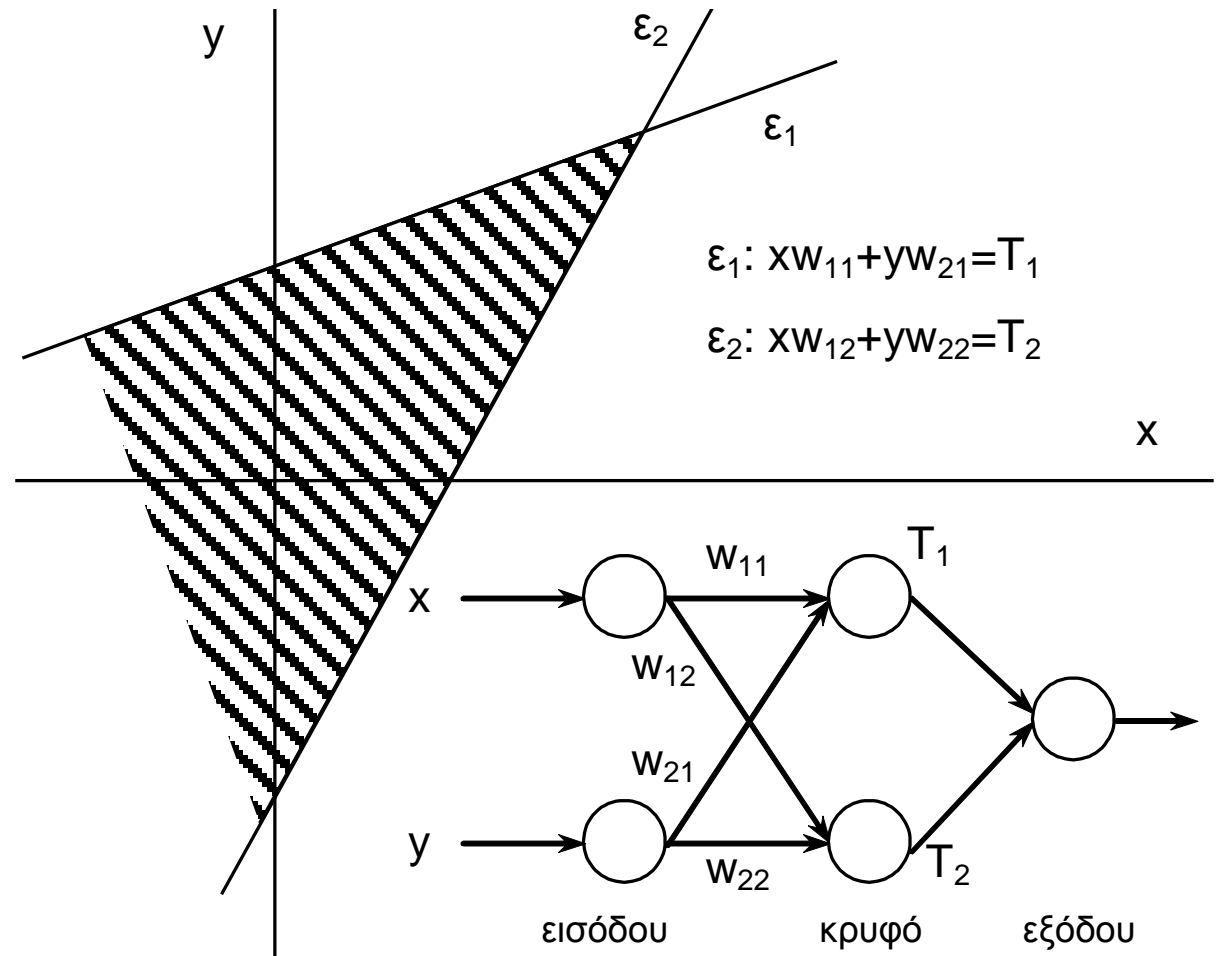
## Συνάρτηση XOR

(μη γραμμικώς διαχωρίσιμη)

Τέτοια προβλήματα δεν μπορούν να μοντελοποιηθούν με perceptron και 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.
  - ❑ Παράγονται περισσότερες ευθείες.
- ❖ Στο σχήμα:
  - ❑ Οι δύο νευρώνες του κρυφού επιπέδου παράγουν τις ευθείες  $\epsilon_1$  και  $\epsilon_2$ .
  - ❑ Η έξοδος πρέπει να είναι στην γραμμοσκιασμένη περιοχή.
  - ❑ Αν υπήρχε και τρίτος νευρώνας θα υπήρχε και  $\epsilon_3$  που θα βελτίωνε περισσότερο την διαχωρισιμότητα.



## Κανόνας Δέλτα (1/2)

- Αποτελεί γενίκευση του αλγορίθμου εκπαίδευσης του perceptron.
  - Προκύπτει θεωρητικά από την ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος (mean square error) των διανυσμάτων εκπαίδευσης.
  - Δε μπορεί να εφαρμοστεί σε δίκτυα με κρυφά επίπεδα.
    - ✓ Επειδή δεν είναι γνωστή η επιθυμητή έξοδος για τους νευρώνες των κρυφών επιπέδων.
- Μέσο τετραγωνικό σφάλμα  $E$  στο στοιχειώδες perceptron, για  $p$  διανύσματα εκπαίδευσης:
  - $input_k$ : σήμα εισόδου του νευρώνα (χωρίς συνάρτηση ενεργοποίησης)
  - $t_k$ : επιθυμητή έξοδος
- Συνολικό σήμα εισόδου για κάποιο διάνυσμα εκπαίδευσης  $p$ :
- Ο κανόνας Δέλτα ακολουθεί την αρνητική κλίση της επιφάνειας σφάλματος, με κατεύθυνση προς το ελάχιστό της. Δηλαδή:
- Η παράγωγος του  $E$  ως προς τα  $w$  είναι:
- Προκύπτει ότι η μεταβολή στην τιμή του βάρους  $w_i$ , εξαιτίας της εκπαίδευσης με ένα μόνο από τα διανύσματα εκπαίδευσης, δίνεται από τη σχέση:  $\Delta w_i = w_{i(new)} - w_{i(old)} = d(t - input)x_i$
- $input$  το συνολικό σήμα εισόδου του νευρώνα,  $t$  είναι η επιθυμητή έξοδος,  $w_{i(old)}$  και  $w_{i(new)}$  η παλιά και η νέα τιμή του βάρους  $w_i$  αντίστοιχα,  $x_i$  η επιμέρους είσοδος  $i$  της οποίας το βάρος αναπροσαρμόζουμε και  $d$  ο ρυθμός μάθησης (learning rate).

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

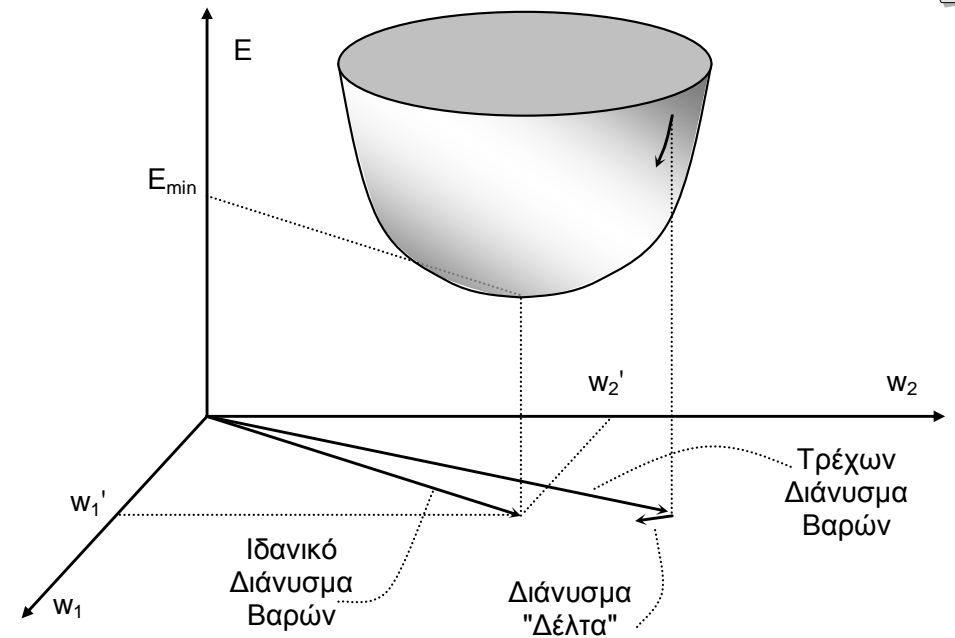
$$input = \sum_{i=1}^n w_i x_i$$

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

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

## Κανόνας Δέλτα (2/2)

- ❖ Γεωμετρική αναπαράσταση κανόνα Δέλτα για στοιχειώδες perceptron με 2 εισόδους.
- ❖ Ονομάζεται και κανόνας της επικλινούς καθόδου (*gradient descent rule*) εξαιτίας του ότι ακολουθεί την αρνητική κλίση της επιφάνειας σφάλματος, με κατεύθυνση προς το ελάχιστό της.



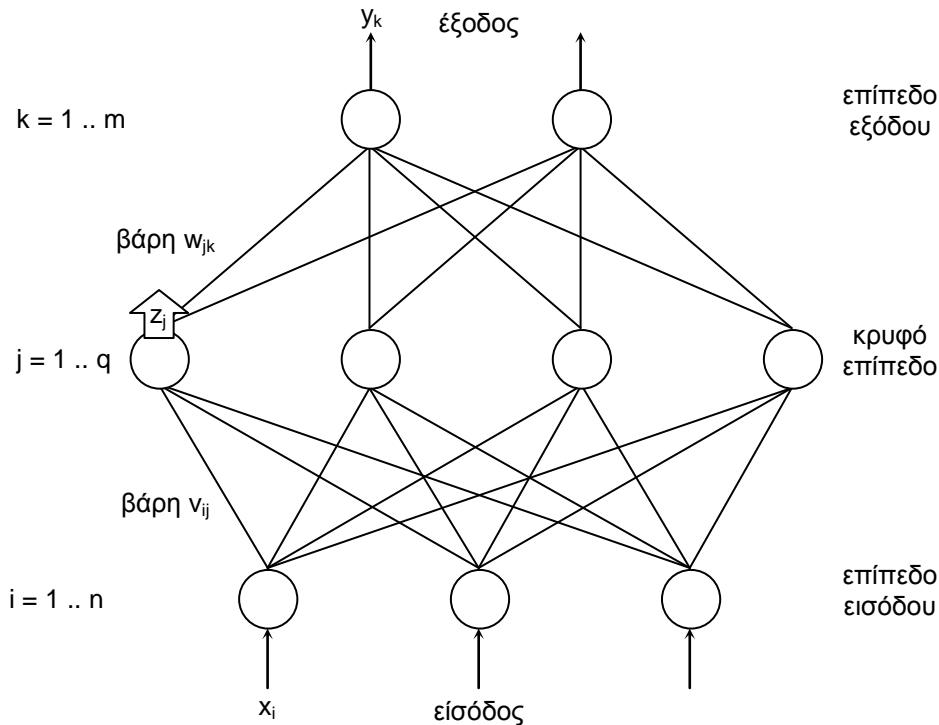
$$\Delta w_i = w_{i(new)} - w_{i(old)} = d(t - input) x_i$$

- ❑ ισχύει και στην περίπτωση που αντί του συνολικού σήματος εισόδου *input*, χρησιμοποιηθεί η πραγματική έξοδος *y* του νευρώνα: 
$$y = f\left(\sum_{i=0}^n w_i x_i\right)$$
- ❑ Αν η *f* είναι μια συνεχής συνάρτηση ενεργοποίησης τότε η σχέση για το  $\Delta w_i$  περιέχει και παράγοντα *f'* (1<sup>η</sup> παράγωγος) ο οποίος πριν εκφυλιζόταν).
- ❖ Ο αλγ. Δέλτα δεν μπορεί να εφαρμοστεί αυτούσιος σε δίκτυα με κρυφά επίπεδα.
  - ❑ δεν είναι γνωστή με ακρίβεια η επιθυμητή έξοδος *t* στους νευρώνες των κρυφών επιπέδων
- ❖ Αποτέλεσμα: Η έρευνα στα ΤΝΔ περιορίστηκε σημαντικά για πολλά χρόνια.
- ❖ Η λύση δόθηκε με τη μέθοδο της ανάστροφης μετάδοσης του λάθους
  - ❑ μαθηματικός τρόπος μεταφοράς του σφάλματος του επιπέδου εξόδου προς τα πίσω



# Ανάστροφη Μετάδοση Λάθους (*back propagation*)

- ❖ Αποτελεί την πιο γνωστή μέθοδο εκπαίδευσης ΤΝΔ πολλών επιπέδων.
- ❖ 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.
  - ❑ Γίνεται έτσι δυνατό να υπολογίζονται οι διορθώσεις στα βάρη κάθε νευρώνα ξεχωριστά.



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

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

- ❖ Αποδεικνύεται ότι (σχέσεις ΓΚΔ):

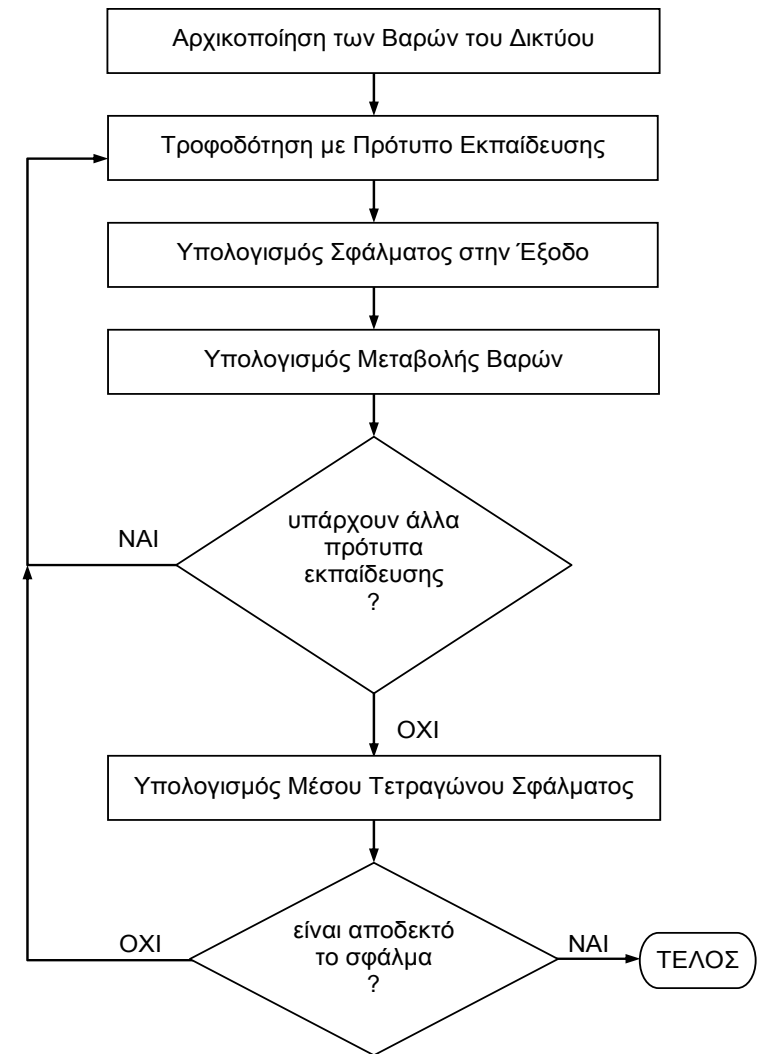
- ❑ για επίπεδο εξόδου:  $\Delta w_{jk} = d \cdot \delta_k \cdot z_j$  με  $\delta_k = (t_k - y_k) f'(input_k)$

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

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

- ❖ Η αναπροσαρμογή των βαρών γίνεται από το επίπεδο εξόδου προς το εισόδου.
  - ❑ *ανάστροφο πέρασμα (backward pass)* ή *ανάστροφη μετάδοση (back propagation)*

- ❖ Είναι μία 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.
- ❖ Αναζήτηση τύπου αναρρίχησης λόφου:
  - ❑ Η διόρθωση που γίνεται κάθε φορά προσπαθεί να ελαχιστοποιήσει το σφάλμα κάνοντας εκείνες τις αλλαγές που φαίνεται να το μειώνουν τοπικά.
- ❖ Προβλήματα:
  - ❑ Τοπικά ελάχιστα: η εγγενής αδυναμία της αναζήτησης αναρρίχησης λόφων να βρει το ολικό ελάχιστο
  - ❑ *Network paralysis*: ένα ή περισσότερα βάρη έχουν σταθερά υψηλές απόλυτες τιμές και δεν τροποποιούνται σημαντικά σε κάθε διόρθωση.
- ❖ Τα διανύσματα εκπαίδευσης δεν πρέπει να μεταβάλλονται: δεν επέρχεται σύγκλιση.
  - ❑ σε αντίθεση με τα βιολογικά συστήματα που μαθαίνουν ακόμη κι όταν αλλάζουν τα δεδομένα που τους παρουσιάζονται.
- ❖ 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.
  - ❑ Η ανάκληση πληροφορίας γίνεται με βάση το βαθμό ομοιότητας μεταξύ του διάνυσματος εισόδου και των αποθηκευμένων προτύπων και όχι κάποια διεύθυνση μνήμης.
- ❖ Ανάλογα με τις διαφορές μεταξύ εισόδου και εξόδου διακρίνουμε:
  - ❑ *αυτοσυσχετιζόμενες μνήμες (auto-associative memories)*: ανακαλούν αποθηκευμένα πρότυπα που μοιάζουν κατά το δυνατό σε δοθέν διάνυσμα εισόδου
  - ❑ *ετεροσυσχετιζόμενες μνήμες (hetero-associative memories)*: είσοδος και έξοδος διαφέρουν, όχι μόνο στο περιεχόμενο αλλά ενδεχομένως και στη δομή.
- ❖ Ανάλογα με το αν η έξοδός τους μπορεί να είναι προϊόν παρεμβολής, διακρίνουμε:
  - ❑ *με δυνατότητα παρεμβολής (interpolative associative memories)*: αν π.χ. συσχετίζουν το 1 με το 5 και το 2 με το 6, τότε σε περίπτωση εισόδου 1.4 θα παράγουν έξοδο 5.4.
  - ❑ *προσαυξητική μνήμη συσχέτισης (accretive associative memory)*: στο προηγούμενο παράδειγμα θα έδινε αποτέλεσμα 5 (γιατί το 1.4 είναι πλησιέστερα στο 1).
- ❖ Τύποι ΤΝΔ που συνιστούν μνήμες συσχέτισης:
  - ❑ Γραμμικοί Συσχετιστές
  - ❑ Δίκτυα Hopfield
  - ❑ Μνήμες Συσχέτισης Διπλής Κατεύθυνσης

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

- ❖ ΤΝΔ πρόσθιας τροφοδότησης: η έξοδος παράγεται με ένα μόνο πρόσθιο πέρασμα.
- ❖ Οι νευρώνες επιπέδου εξόδου: γραμμική συνάρτηση ενεργοποίησης
- ❖ Πρότυπα συσχέτισης: δυαδικά (binary, 0 ή/και 1) ή διπολικά (bipolar, +1 ή/και -1)
- ❖ Εκπαίδευση:

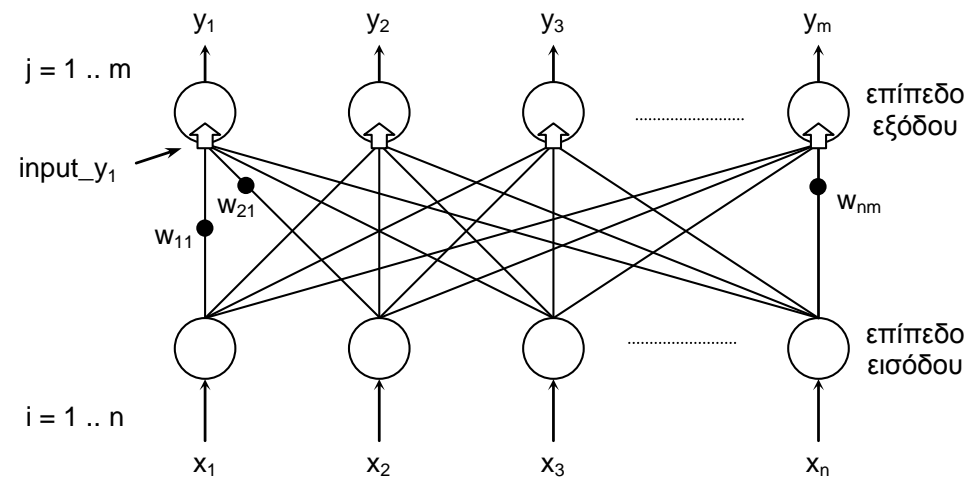
- ❑  $\Delta w_{ij} = x_i \cdot y_j$  (κανόνας Hebb)
- ❑ μετά από  $p$  ζευγάρια εκπαίδευσης

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

- ❑ 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

$$y_j = \begin{cases} 1 & , y_{input\_j} > 0 \\ 0 & , y_{input\_j} \leq 0 \end{cases} \quad \begin{matrix} \Leftarrow \text{για δυαδικά διανύσματα } X \text{ και } Y \\ \text{για διπολικά διανύσματα } X \text{ και } Y \Rightarrow \end{matrix} \quad y_j = \begin{cases} +1 & , y_{input\_j} \geq 0 \\ -1 & , y_{input\_j} < 0 \end{cases}$$

□  $y_{input\_j}$  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)*.

- αποτέλεσμα: η έξοδος θα είναι πάντα ίδια με την είσοδο, η οποία βέβαια δεν θα είναι πάντα κάποιο από τα πρότυπα που αρχικά απομνημόνευσε το δίκτυο.

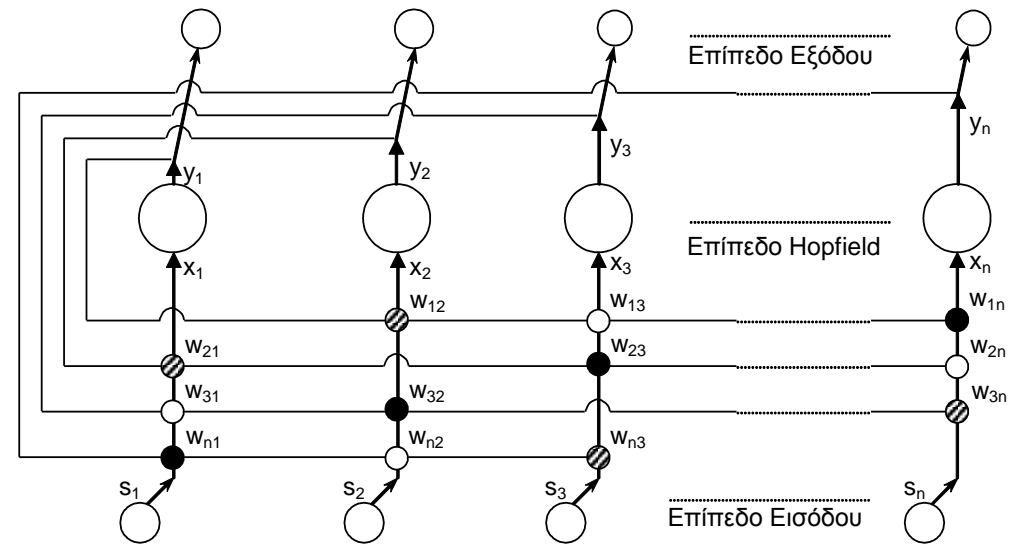
- ❖ Η **τέλεια ανάκληση** είναι δυνατή μόνο όταν τα διανύσματα εισόδου είναι ορθογώνια μεταξύ τους, δηλαδή όταν  $X_i \cdot X_j = 0$  για  $i \neq j$ .

- μόνο τότε ο όρος αλληλεπίδρασης (cross-talk) της συσχέτισης είναι μηδενικός

- ❖ Χωρητικότητα (με τέλεια ανάκληση) δομής  $n \times n$ :  $n-1$  διπολικά ζευγάρια

# Δίκτυα Hopfield (1/3)

- ❖ Αποτελούνται από ένα μόνο επίπεδο νευρώνων, το επίπεδο Hopfield.
- ❖ Δίκτυα με ανατροφοδότηση (*recurrent*): οι συνδέσεις μεταξύ των νευρώνων είναι αμφίδρομες (επιπλέον, έχουν και συμμετρικά βάρη).
- ❖ Λειτουργούν ως αυτοσυσχετιζόμενες μνήμες αλλά 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  $w_{ii}=0$  (οι νευρώνες δε συνδέονται με τον εαυτό τους).
- ❖ Στη συνάρτηση ενεργοποίησης,  $y_i'$  είναι η έξοδος του προηγούμενου κύκλου υπολογισμών, ενώ το κατώφλι  $\theta_i$  συνήθως τίθεται ίσο με μηδέν.

# Δίκτυα Hopfield (2/3)

❖ Παράδειγμα: θέλουμε να αποθηκευτεί το διάνυσμα (πρότυπο)  $S_1=(1, 1, 1, 0)$

❖ 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  $S_1$  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.

- ❑ μπορεί να ανακαλέσει κάποιο πρότυπο ακόμα και με αλλοιωμένη είσοδο

❖ 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  $j$

β) Υπολογίζεται η είσοδος του 
$$x_j = \sum_{i=1, i \neq j}^n y_i w_{ij} + s_j$$

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  $W_I$  (προηγούμενη διαφάνεια)
- ❖ Τροφοδοτείται με το διάνυσμα  $S'=(s'_1, s'_2, s'_3, s'_4)=(0, 0, 1, 0)$ , αλλοιωμένη μορφή του  $S_I$  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**

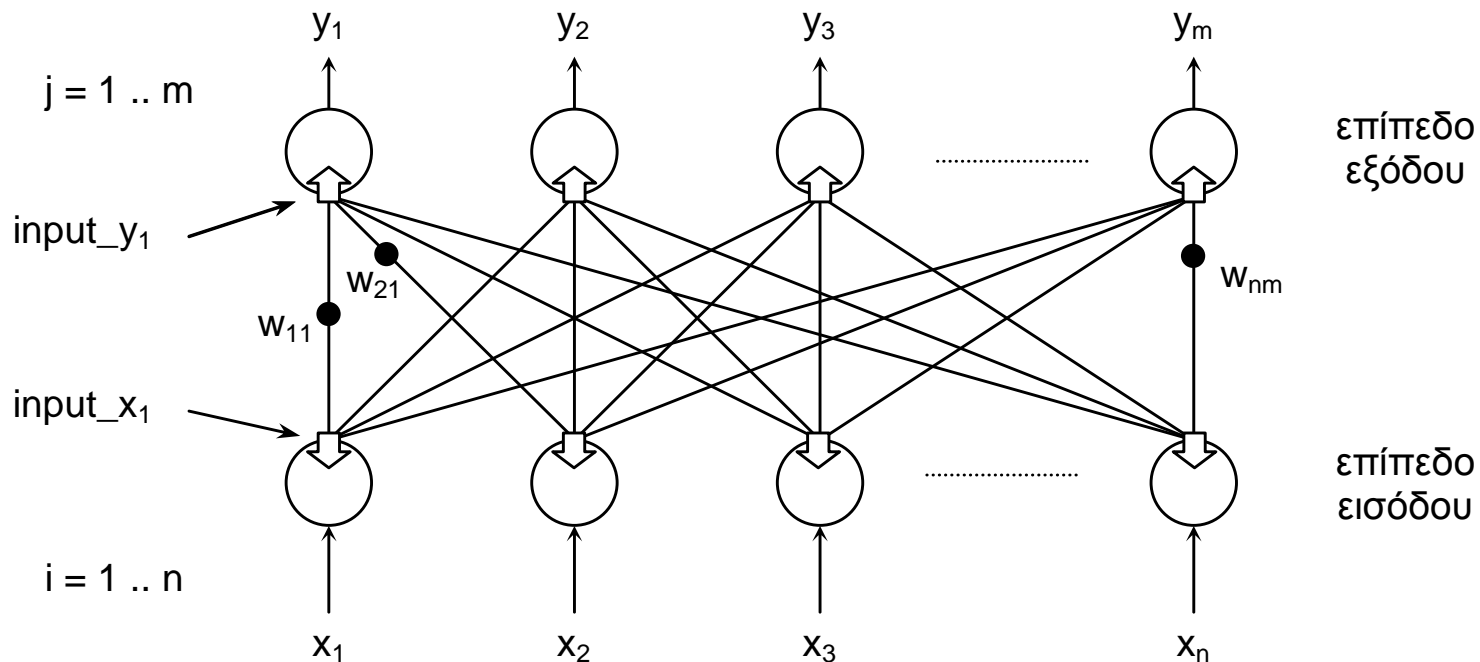
$$y_i = \begin{cases} 1 & , x_i > \theta_i \\ y_i' & , x_i = \theta_i \\ -1 & , x_i < \theta_i \end{cases}$$

**Σημείωση:** στα  $\Sigma$  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 εισόδου  $X$  και εξόδου  $Y$ : διπολικά ή δυαδικά.

# Μνήμες Συσχέτισης Διπλής Κατεύθυνσης

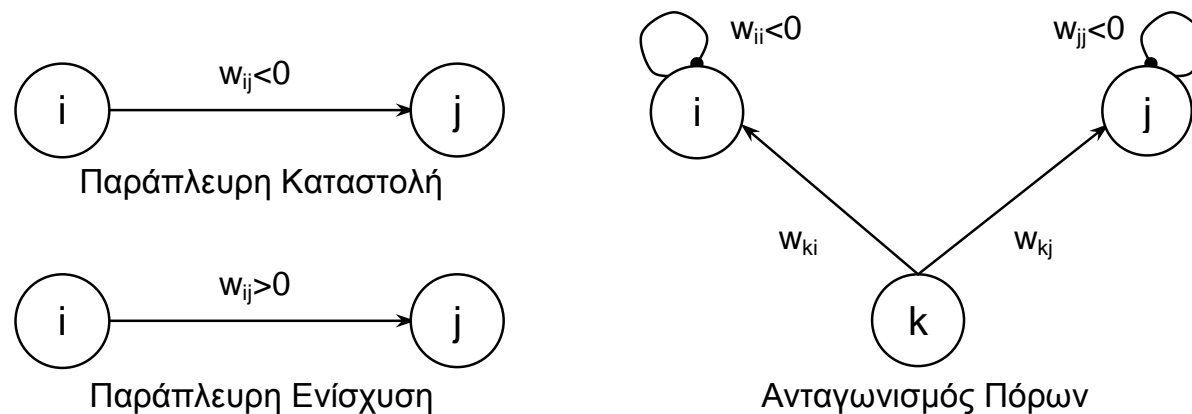
- ❖ Εκπαίδευση: με ζευγάρια διπολικών διανυσμάτων εισόδου-εξόδου και βηματική συνάρτηση ενεργοποίησης.

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

- ❑ όπου  $y_j'$  είναι η έξοδος του προηγούμενου κύκλου υπολογισμών
- ❑ Αντίστοιχες σχέσεις ισχύουν για τους νευρώνες του επιπέδου  $X$
- ❖ 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
- ❖ Η χωρητικότητα του μοντέλου BAM είναι της τάξης  $O(\max(n, m))$ .

# 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.
  - ❑ δημιουργείται ανταγωνισμός για το 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*): η έξοδος κάθε νευρώνα επηρεάζει αρνητικά ή θετικά τους γειτονικούς μέσω αρνητικών ή θετικών βαρών.
  - ❑ **Ανταγωνισμός πόρων** (*resource competition*): η έξοδος ενός νευρώνα κατανέμεται στους συνδεδεμένους στην έξοδο νευρώνες αναλογικά με τα βάρη των συνδέσεων και την έξοδο που αυτοί παράγουν.
    - βιολογικά αποδεκτό - μοντελοποιεί και το βιολογικό φαινόμενο της *εξασθένισης* (*self decay*).



# 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.
- ❖ Δεν αποτελούν υλοποίηση των παράπλευρων διεγέρσεων και καταστολών αλλά δίνουν παρόμοια αποτελέσματα και υλοποιούνται εύκολα σε H/Y.

# 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

• για δεδομένο διάνυσμα εισόδου  $S$  της μορφής  $S=(s_1, s_2, \dots, s_k)$  η κανονικοποίηση γίνεται μέσω της σχέσης:

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

- $s_i$  και  $s_i'$  είναι η αρχική και η κανονικοποιημένη συνιστώσες για το διάνυσμα εισόδου  $S$
- όμοια κανονικοποιούνται και τα αρχικά βάρη εισόδου  $W_i=(w_{i1}, w_{i2}, \dots, w_{ik})$

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

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

- με χρήση Ευκλείδιας απόστασης:
- ο νευρώνας  $c$  με τη μικρότερη τιμή  $d$  είναι ο "νικητής" και δικαιούται να "εκπαιδευτεί"

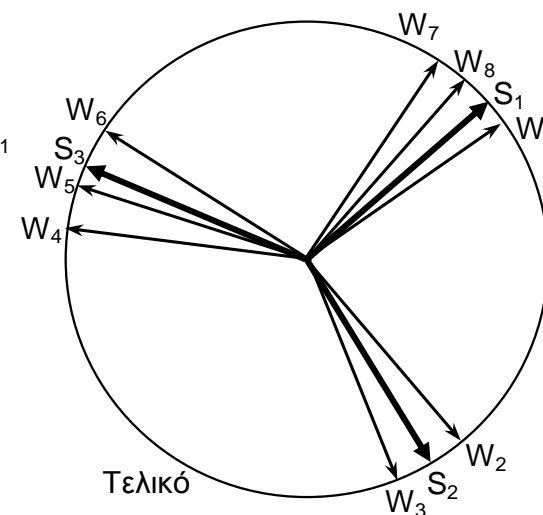
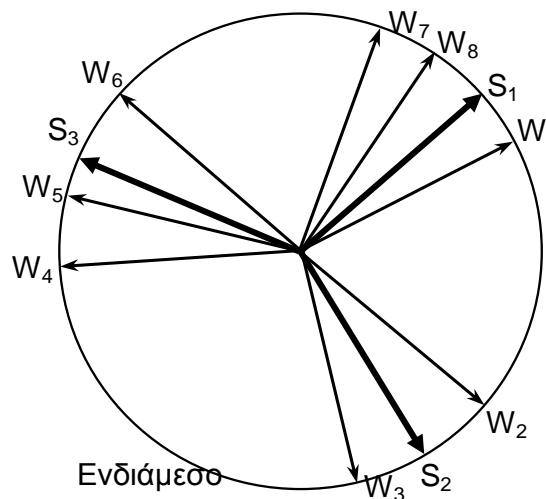
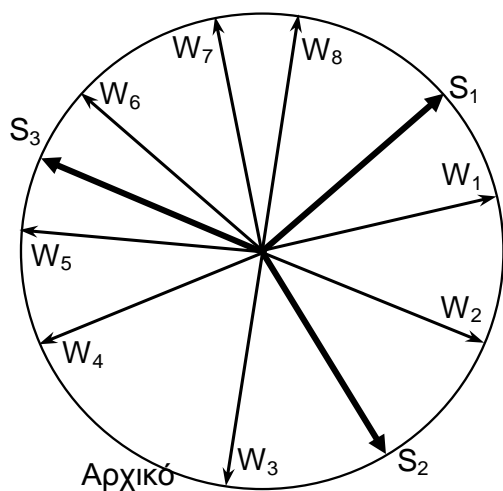
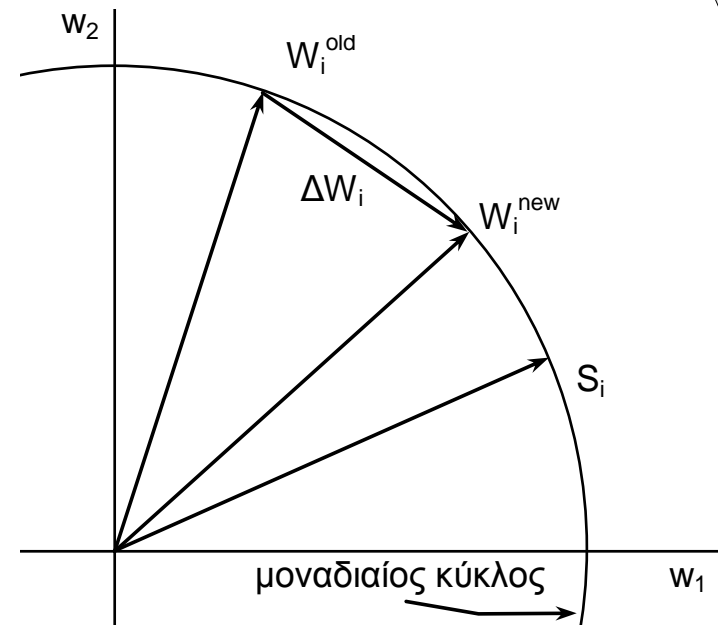
$$\Delta w_{ij} = \begin{cases} n(s_j - w_{ij}) & \text{αν ο νευρώνας } i \text{ ανήκει στη γειτονιά του } c \\ 0 & \text{αν ο νευρώνας } i \text{ δεν ανήκει στη γειτονιά του } c \end{cases}$$

- $n$  είναι ο ρυθμός εκπαίδευσης και  $j=1, \dots, k$ .
- εκπαίδευση δικαιούνται και οι νευρώνες που βρίσκονται μέσα στη γειτονιά του νικητή, που καθορίζεται από πριν (συνήθως βάσει γεωμετρικών χαρακτηριστικών)

❖ Πριν αρχίσει ο επόμενος κύκλος εκπαίδευσης, ο ρυθμός εκπαίδευσης  $n$  και το μέγεθος της γειτονίας ελαττώνονται.

# Δίκτυα Kohonen (3/5)

- ❖ Οπτικοποίηση ενός βήματος εκπαίδευσης σε δίκτυο Kohonen με 2 στοιχεία εισόδου:
  - ❑ Μετά από ένα βήμα εκπαίδευσης, το διάνυσμα βαρών  $W_i$  μετατοπίζεται κατά  $\Delta W_i$  και πλησιάζει το διάνυσμα  $S_i$  που αντιστοιχεί στο διάνυσμα εκπαίδευσης.
  - ❑ Όλα τα στοιχεία είναι κανονικοποιημένα.
- ❖ 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



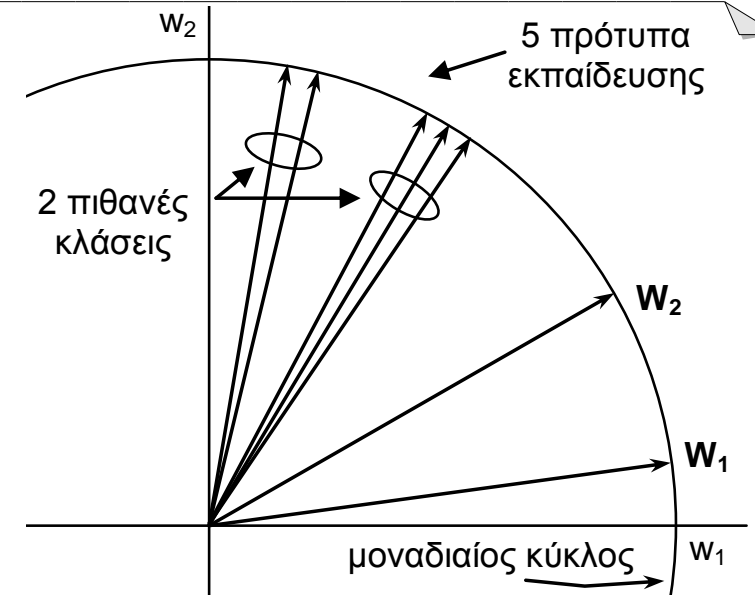
# Δίκτυα Kohonen (4/5)

## ❖ Κακή αρχικοποίηση τιμών για τα βάρη (σχήμα).

- ❑ φαινομενικά υπάρχουν 2 κλάσεις
- ❑ η κλάση που ορίζουν τα 3 διανύσματα εκπαίδευσης θα "τραβήξει" πάνω της και τα δύο διανύσματα βαρών
- ❑ θα έπρεπε τα  $w_1$  και  $w_2$  να αρχικοποιηθούν εκατέρωθεν των δύο φαινομενικών ομάδων

## ❖ 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)

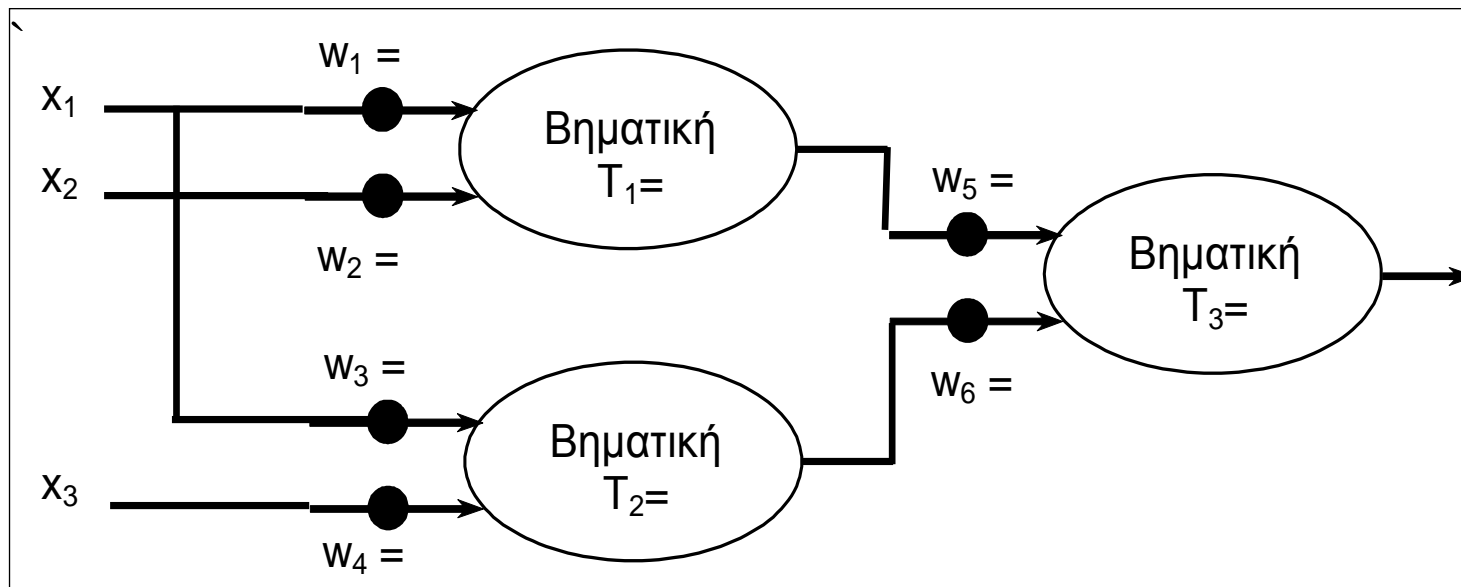
- ❖ Η διαδικασία επαναλαμβάνεται για τα υπόλοιπα διανύσματα εκπαίδευσης.
  - μετά το πέρας ενός κύκλου εκπαίδευσης:  
 $W_1=(0.032, 0.096, 0.680, 0.984)$  και  $W_2=(0.968, 0.304, 0.112, 0.048)$
- ❖ Στη συνέχεια, μεταβάλλεται ο ρυθμός εκπαίδευσης σε 0.3 και αρχίζει νέα εποχή.
  - μετά το πέρας του δεύτερου 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<sup>η</sup> ή 2<sup>η</sup>) στην οποία ανήκει κάθε διάνυσμα  $S_j$  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<sup>η</sup> κλάση
    - $j=2$   $W_1 \cdot S_2=1$  και  $W_2 \cdot S_2=0$  άρα το  $S_2$  ανήκει στην 1<sup>η</sup> κλάση
    - $j=3$   $W_1 \cdot S_3=0$  και  $W_2 \cdot S_3=1$  άρα το  $S_3$  ανήκει στην 2<sup>η</sup> κλάση
    - $j=4$   $W_1 \cdot S_4=1.5$  και  $W_2 \cdot S_4=0$  άρα το  $S_4$  ανήκει στην 1<sup>η</sup> κλάση

# Εφαρμογές Νευρωνικών Δικτύων

- ❖ Δημοφιλή σε προβλήματα που περιέχουν μη-προβλέψιμες λειτουργίες και τα οποία δεν είναι πλήρως κατανοητά.
- ❖ Κατηγοριοποίηση (ταξινόμηση)
  - ✓ Ιατρική: κατηγοριοποίηση (ως πρώτη διάγνωση) ιατρικών εικόνων από εξετάσεις υπέρηχων, τεστ ΠΑΠ, κτλ.
  - ✓ Άμυνα: κατηγοριοποίηση εικόνων προερχόμενων από συσκευές όπως radar, sonar, κτλ.
  - ✓ Γεωργία: έλεγχος καλλιεργειών σε συνδυασμό με δορυφορικά συστήματα τηλε-επισκόπησης.
  - ✓ Οικονομία/επιχειρήσεις: κατηγοριοποίηση πελατών με βάση τις αγοραστικές τους συνήθειες.
- ❖ Αναγνώριση
  - ✓ Τράπεζες: γνησιότητα υπογραφής και τραπεζογραμματίων.
  - ✓ Πληροφορική και τηλεπικοινωνίες: αναγνώριση ήχου, εικόνας και γραπτού κειμένου (OCR), κτλ.
- ❖ Αποτίμηση
  - ✓ Άμυνα: παρακολούθηση στόχων.
  - ✓ Ασφάλεια: εντοπισμός κίνησης (*motion detection*), ταύτιση δακτυλικών αποτυπωμάτων, ανάλυση εικόνας σε συστήματα επιτήρησης.
  - ✓ Μηχανολογία: παρακολούθηση, επιθεώρηση και έλεγχος προϊόντων.
- ❖ Πρόβλεψη
  - ✓ Οικονομία/επιχειρήσεις: πρόβλεψη ισοτιμίας νομισμάτων και τιμών μετοχών (συνήθως βραχυπρόθεσμη), πρόβλεψη πωλήσεων, κτλ.
  - ✓ Γεωργία: πρόβλεψη παραγωγής, κυρίως με χρήση δορυφορικών εικόνων.
  - ✓ Μετεωρολογία: πρόβλεψη καιρού.

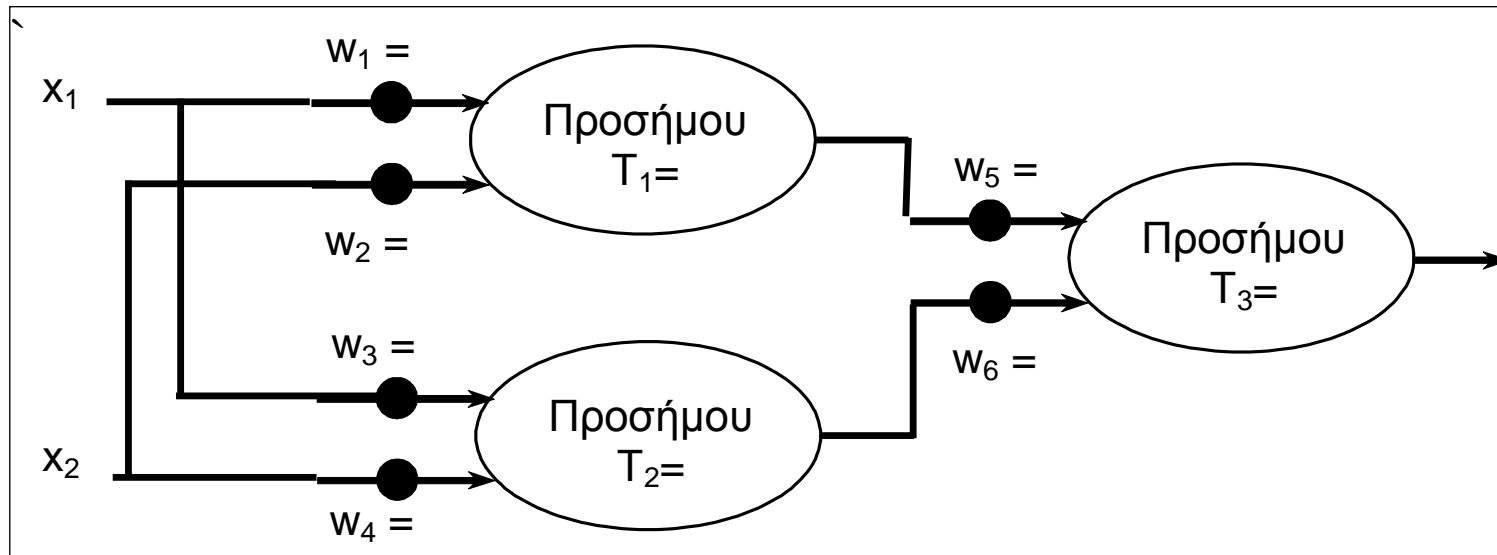
# Ασκήσεις

- ❖ 11.1 Στο παρακάτω σχήμα δίνεται ένα νευρωνικό δίκτυο με 3 εισόδους και τρεις νευρώνες. Συμπληρώστε τις τιμές των βαρών και τα κατώφλια των τριών βηματικών συναρτήσεων, ώστε το δίκτυο να βγάζει ως αποτέλεσμα 1 αν όλες οι εισοδοί έχουν τιμή 1 και αποτέλεσμα 0 σε κάθε άλλη περίπτωση.



- ❖  $w_1 = \dots, w_2 = \dots, w_3 = \dots, w_4 = \dots, w_5 = \dots, w_6 = \dots, T_1 = \dots, T_2 = \dots, T_3 = \dots$

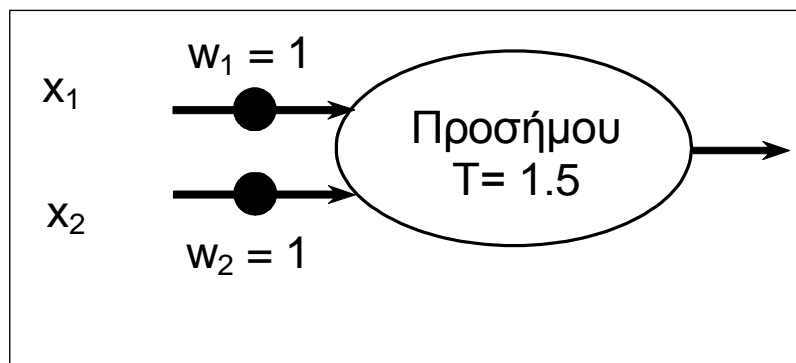
- ❖ 11.2 Στο παρακάτω σχήμα δίνεται ένα νευρωνικό δίκτυο με 2 εισόδους και τρεις νευρώνες. Συμπληρώστε τις τιμές των βαρών και τα κατώφλια των τριών συναρτήσεων προσήμου, ώστε το δίκτυο να λειτουργεί όπως η συνάρτηση XOR.



- ❖  $w_1 = \dots, w_2 = \dots, w_3 = \dots, w_4 = \dots, w_5 = \dots, w_6 = \dots, T_1 = \dots, T_2 = \dots, T_3 = \dots$

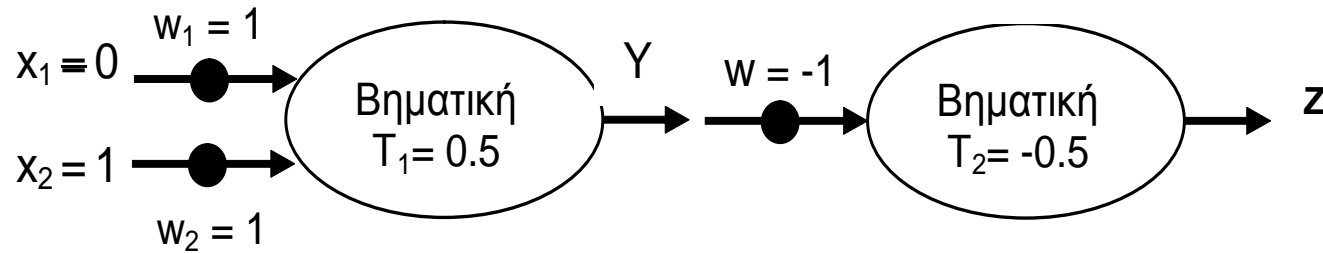


- ❖ 11.3 Στο παρακάτω σχήμα δίνεται ένας νευρώνας με 2 εισόδους με βάρη  $w_1=1$  και  $w_2=1$  και Συνάρτηση Ενεργοποίησης τη συνάρτηση προσήμου με κατώφλι  $T=1.5$ . Να υπολογίσετε την έξοδο του όταν η είσοδος είναι:
- α)  $x_1=0, x_2=1,$   
β)  $x_1=1$  και  $x_2=1.$

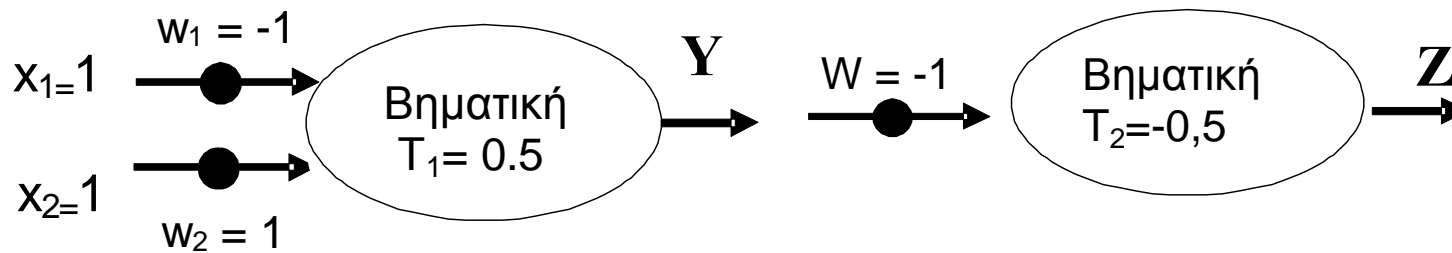




❖ 11.4. Ποια είναι ή έξοδος  $Y$  του συστήματος για τη δεδομένη είσοδο.



❖ 11.5. Συμπληρώστε τις τιμές των  $Y$  και  $Z$ .



❖ 11.6 Να σχεδιαστεί τεχνητός νευρώνας που να υλοποιεί τη λογική συνάρτηση NOR. Η συνάρτηση αυτή επιστρέφει τιμή 1 όταν και οι δύο είσοδοι είναι 0 και τιμή 0 σε κάθε άλλη περίπτωση. Για τον νευρώνα να ορίσετε τη συνάρτηση κατωφλίου και τα βάρη.