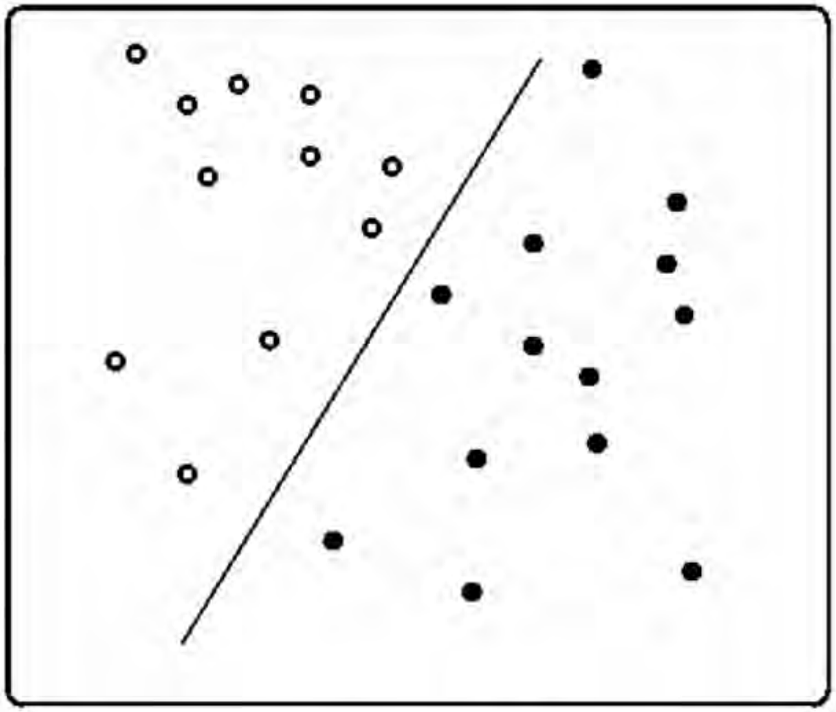
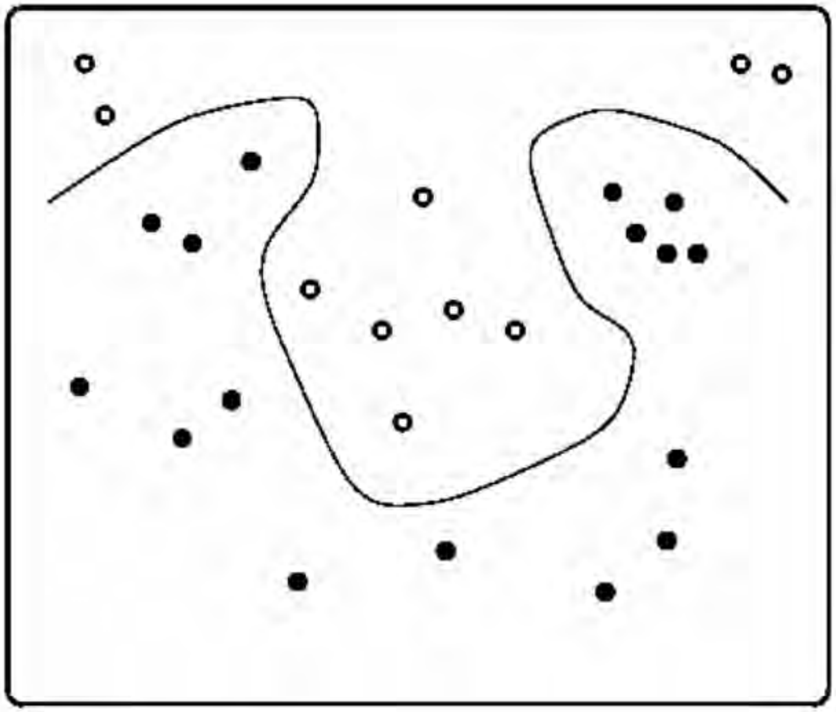
Γραμμικός ταξινομητής

Σύστημα ταξινόμησης με χρήση γραμμικών διακριτικών συναρτήσεων





$$D(\mathbf{x}) = w_1 x_1 + w_2 x_2 + \ldots + w_\nu x_\nu + \ldots + w_N x_N + w_{N+1} = \sum_{\nu=1}^{N} w_\nu x_\nu + w_{N+1}$$

$$D(\mathbf{x}) > 0 \Rightarrow \mathbf{x} \in C_1,$$

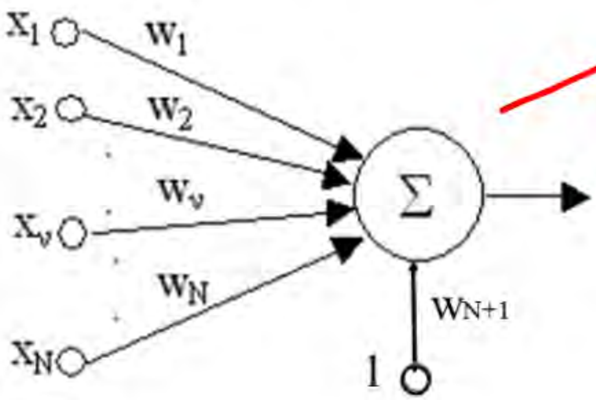
$$D(\mathbf{x}) < 0 \Rightarrow \mathbf{x} \in C_2,$$

$$D(\mathbf{x}) = 0 \Rightarrow \mathbf{x} \in (\varepsilon)$$

Γραμμικός ταξινομητής

Σύστημα ταξινόμησης με χρήση γραμμικών διακριτικών συναρτήσεων

$$D(\mathbf{x}) = \sum_{\nu=1}^{N} w_{\nu} x_{\nu} + w_{N+1}$$



Νευρώνας

Οι συντελεστές w αντιστοιχούν στις συνάψεις

Συστήματα διασυνδεδεμένων νευρώνων συνθέτουν τα *Τεχνητά νευρωνικά δίκτυα*
Artificial Neural Networks

Νευρωνικά δίκτυα που περιγράφουν γραμμικούς ταξινομητές ονομάζονται *perceptrons*

Τεχνητά νευρωνικά δίκτυα

- Μοντέλα των βασικών στοιχείων του εγκεφάλου και του νευρικού συστήματος

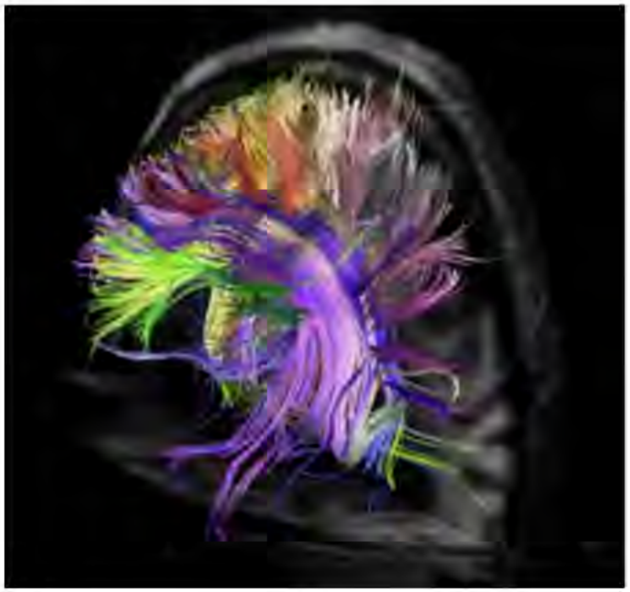
- Υψηλός δείκτης παραλληλίας
  - Επεξεργασία πληροφοριών σαν τον εγκέφαλο και όχι σαν τον υπολογιστή (σειριακά)

- Ικανότητα μάθησης και γενίκευσης
  - Κατανόηση της λειτουργίας τους χωρίς εξωτερική επέμβαση
  - Κατανόηση της λειτουργίας τους μέσω παραδειγμάτων
  - Δίνουν ικανοποιητικό αποτέλεσμα ακόμη και για άγνωστα δεδομένα

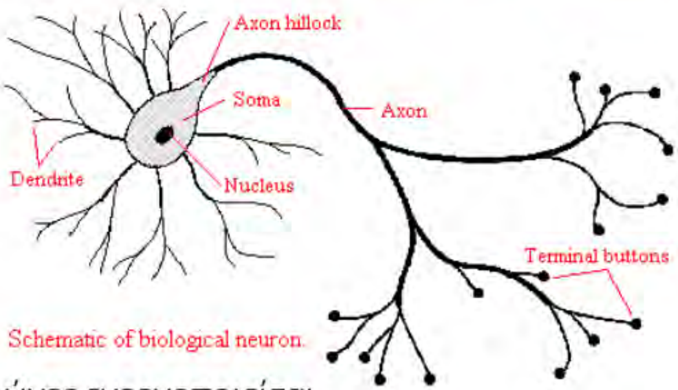
- Βασίζονται σε πολύ απλές αρχές αλλά παρουσιάζουν πολύ πολύπλοκη συμπεριφορά

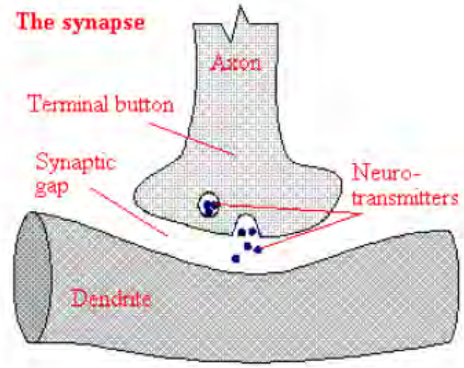
- Εφαρμογές
  - Ισχυρά εργαλεία επίλυσης προβλημάτων
  - Βιολογικά μοντέλα

Βιολογικά νευρωνικά δίκτυα



- Ο εγκέφαλος αποτελείται από περίπου 10 δις διασυνδεδεμένους νευρώνες
  - νεότερες έρευνες θεωρούν λιγότερους
- Κάθε νευρώνας είναι ένα κύτταρο που χρησιμοποιεί βιοχημικές αντιδράσεις για να συλλέξει, να επεξεργαστεί και να μεταδώσει πληροφορίες.



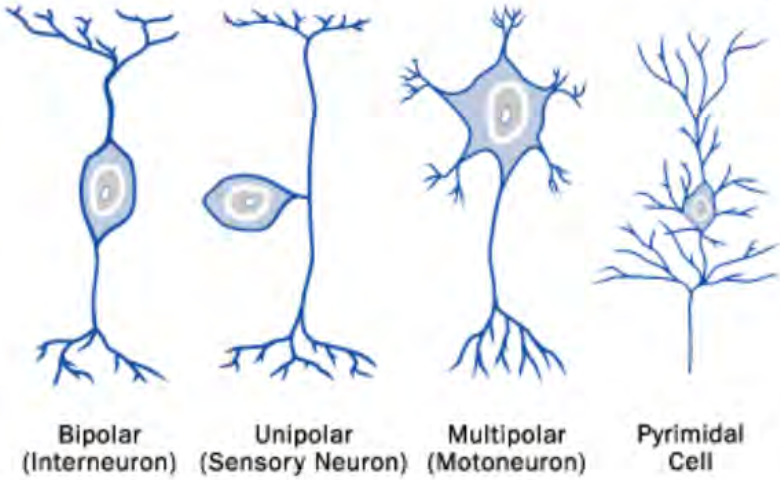


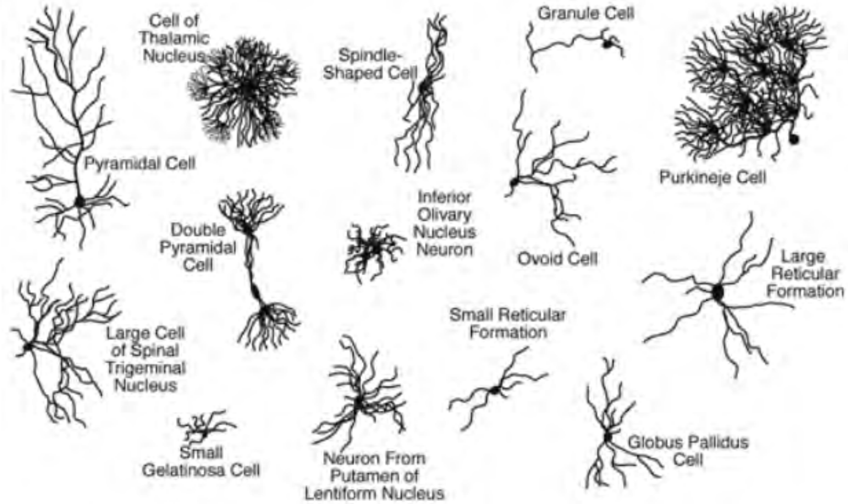
Όταν ένας από τους διασυνδεδεμένους νευρώνες ενεργοποιείται
- ένα θετικό ή αρνητικό φορτίο λαμβάνεται από κάποιον δεντρίτη
- η ένταση όλων των λαμβανόμενων φορτίων αθροίζονται μέσω χωρο-χρονικής άθροισης
- η διαταραχή οδηγείται στο σώμα του νευρώνα
  - όταν η διαταραχή υπερβαίνει το κατώφλι τιμής στο axon hillock το κύτταρο ενεργοποιείται και αποστέλλει πληροφορία μέσω του άξονα με σταθερή ένταση σε κάθε terminal button και από εκεί σε κάθε διασυνδεδεμένο νευρώνα.

Τα φυσικά και νευρωχημικά χαρακτηριστικά κάθε σύναψης καθορίζουν την ένταση και την πολικότητα του σήματος.
Εδώ έγκειται η μεγάλη ευελιξία του εγκεφάλου αλλά και η αδυναμία του.

Το αέριο νεύρων sarin μπορεί να σκοτώσει γιατί αδρανοποιεί ένα χημικό (acetylcholinesterase) που είναι κανονικά υπεύθυνο για την καταστροφή ενός νευρωδιαβιβαστή (acetylcholine). Αυτο σημαίνει ότι άπαξ και ένας νευρώνας ενεργοποιηθεί εξακολουθεί να ενεργοποιεί όλους τους γειτονικούς νευρώνες. Έτσι χάνεται ο έλεγχος των μυών και επέρχεται ασφυξία.

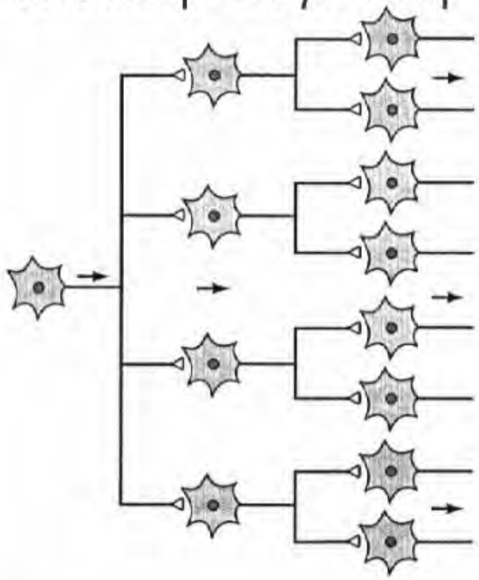
Βιολογικά νευρωνικά δίκτυα

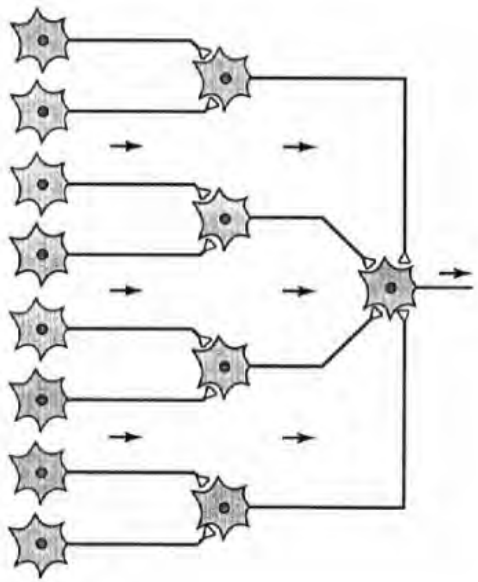




Αρχές οργάνωσης

- Απόκλιση - Σύγκλιση





Βιολογικά νευρωνικά δίκτυα

Περιστέρια ως ειδικοί τέχνης (Watanabe et al. 1995)

Πείραμα:
- Περιστέρια σε Skinner box
- Παρουσίαση πινάκων δύο καλλιτεχνών
  - (π.χ. Chagall / Van Gogh)
- Επιβράβευση για ράμφισμα κατά την παρουσίαση συγκεκριμένου καλλιτέχνη
  - (π.χ. Van Gogh)

Van Gogh - La Roubine du Roi

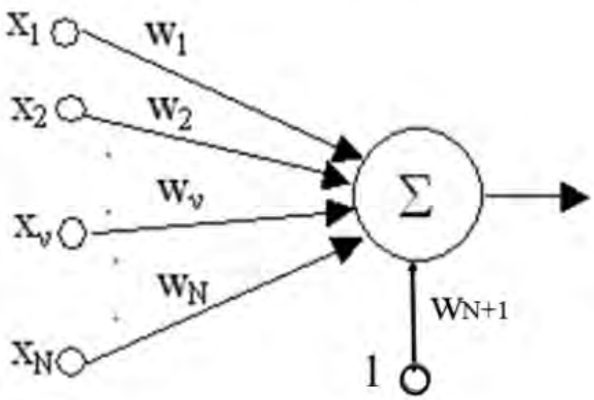
Chagall - The russian wedding

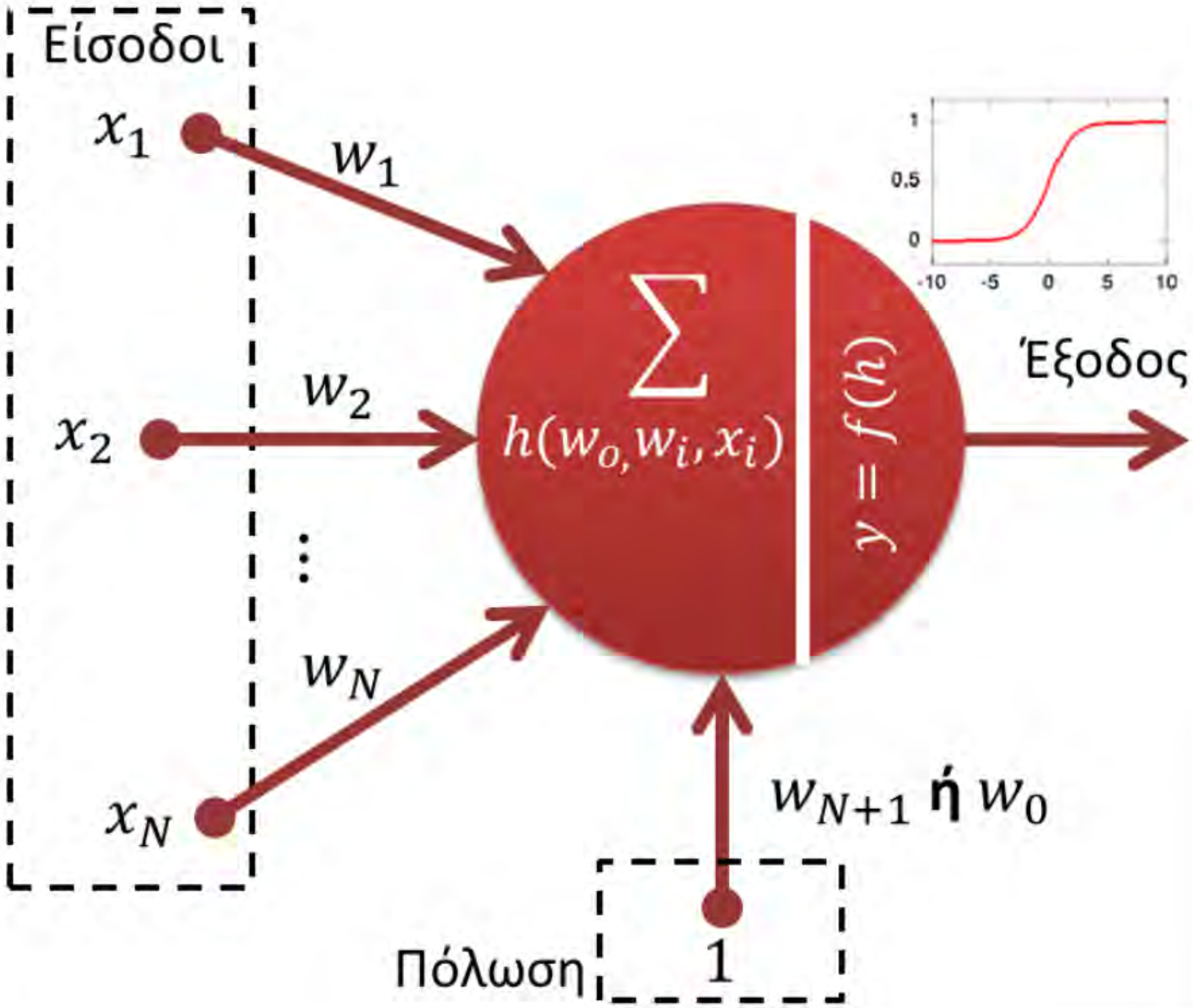
Αποτέλεσμα:
- 95% επιτυχία στα δείγματα εκπαίδευσης
- 85% επιτυχία σε νέα δείγματα

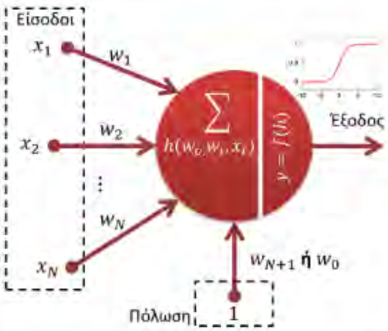
Συμπέρασμα:
- Όχι απλά απομνημόνευση αλλά αναγνώριση και γενίκευση

Μοντελοποίηση βιολογικών νευρώνων

$$D(\mathbf{x}) = \sum_{\nu=1}^{N} w_\nu x_\nu + w_{N+1}$$





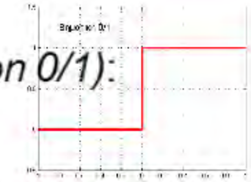


Τεχνητά νευρωνικά δίκτυα

Συναρτήσεις ενεργοποίησης

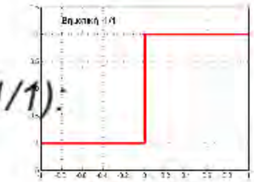
Βηματική 0/1 (step function 0/1):

$$f(\sigma)=\begin{cases}0 & \alpha\nu \quad \sigma\le 0\\ 1 & \alpha\nu \quad \sigma>0\end{cases}$$



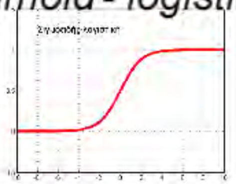
Βηματική -1/1 (step function -1/1):

$$f(\sigma)=\begin{cases}-1 & \alpha\nu \quad \sigma\le 0\\ 1 & \alpha\nu \quad \sigma>0\end{cases}$$



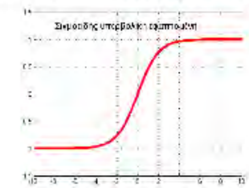
Σιγμοειδής-λογιστική (sigmoid - logistic):

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



Σιγμοειδής-Υπερβολική εφαπτομένη (hyperbolic tangent):

$$f(\sigma)=\tan(\sigma)=\frac{1-e^{-\sigma}}{1+e^{-\sigma}}$$

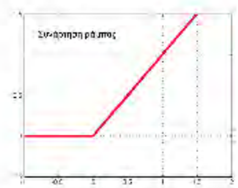


Συνάρτηση κατωφλίου (threshold function):

$$f(\sigma)=\begin{cases}0 & \alpha\nu \quad \sigma\le 0\\ \sigma & \alpha\nu \quad 0<\sigma<1\\ 1 & \alpha\nu \quad 1\le\sigma\end{cases}$$

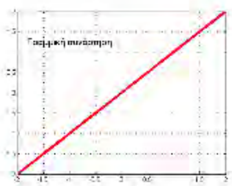
Συνάρτηση ράμπας (ramp function):

$$f(\sigma)=\begin{cases}0 & \alpha\nu \quad \sigma\le 0\\ \sigma & \alpha\nu \quad \sigma>0\end{cases}$$



Γραμμική συνάρτηση (linear function):

$$f(\sigma)=\sigma$$

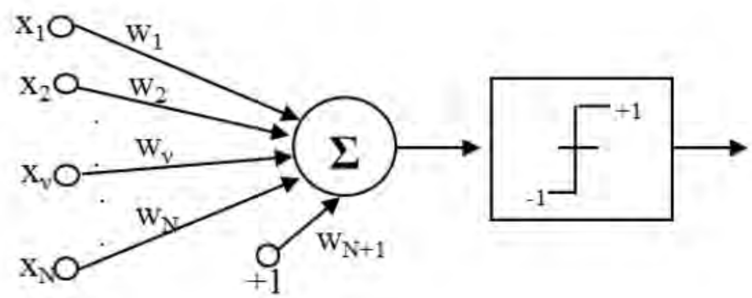


Εκπαίδευση perceptron

Ταξινομητής δύο κλάσεων

Επαναληπτική διαδικασία διόρθωσης σφάλματος:

- αν Α, Β οι δύο κλάσεις
- επιλέγουμε τυχαία ένα πρότυπο από την Α και ένα από τη Β
- φέρουμε μεσοκάθετο υπερεπίπεδο στα τυχαία πρότυπα
- βασική ιδέα της εκμάθησης:
  - διόρθωση του υπερεπιπέδου ώστε να τηρείται το σωστό πρόσημο για όλα τα πρότυπα



$$\boldsymbol{W}_{t+1} = \boldsymbol{W}_t \pm \rho \cdot \widetilde{\boldsymbol{x}}$$

ρ: παράμετρος μάθησης, t: μετρητής επανάληψης

Εκπαίδευση perceptron

Ταξινομητής δύο κλάσεων

Επαναληπτική διαδικασία 6 βημάτων:

1. Επιλογή δύο προτύπων αναφοράς από τις κλάσεις A, B
2. Υπολογισμός μεσοκάθετου υπερεπιπέδου
3. Αρχικοποίηση επαναλήψεων t=0 και θεώρηση:
   $$D(\boldsymbol{X}) = \boldsymbol{W}^T\boldsymbol{X}, \qquad \boldsymbol{W} = [\boldsymbol{w}, w_{N+1}]^T, \boldsymbol{X} = [\boldsymbol{x}, 1]^T$$
4. Υπολογισμός $D_t(\widetilde{\boldsymbol{x}}^A)$ και εκτίμηση προσήμου κάθε κλάσης
5. Επιλογή κάθε προτύπου από το σύνολο εκπαίδευσης και εκτίμηση
   $D_t(\widetilde{\boldsymbol{x}})$ Αν το πρόσημο είναι σωστό $\Rightarrow \boldsymbol{W}_{t+1} = \boldsymbol{W}_t$
   διαφορετικά $\boldsymbol{W}_{t+1} = \boldsymbol{W}_t \pm \rho \cdot \widetilde{\boldsymbol{x}}$
   με πρόσημο ίδιο με την κλάση που ανήκει
6. Αυξάνεται η τιμή του t κατά 1 και η διαδικασία επαναλαμβάνεται από το βήμα 5 για όλα τα πρότυπα του συνόλου εκπαίδευσης

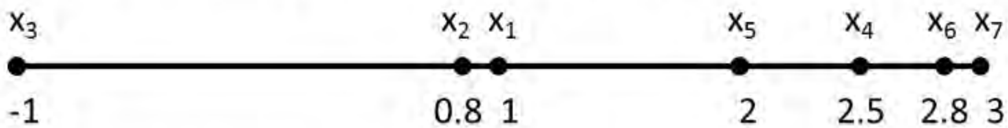
Εκπαίδευση perceptron

Ταξινομητής δύο κλάσεων

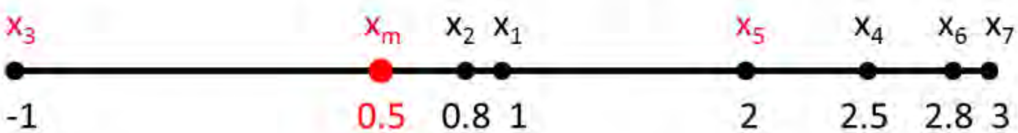
Παράδειγμα σε μία διάσταση

$x^A{}_1 = [1], \quad x^A{}_2 = [0.8], \quad x^A{}_3 = [-1]$

$x^B{}_4 = [2.5], \quad x^B{}_5 = [2], \quad x^B{}_6 = [2.8], \quad x^B{}_7 = [3]$



Επιλέγουμε τυχαία τα $x^A{}_3 = [-1], \quad x^B{}_5 = [2]$



$$x^A{}_3 = [-1], \quad x^B{}_5 = [2] \quad \Rightarrow x_m = \frac{1}{2}$$

αφού:

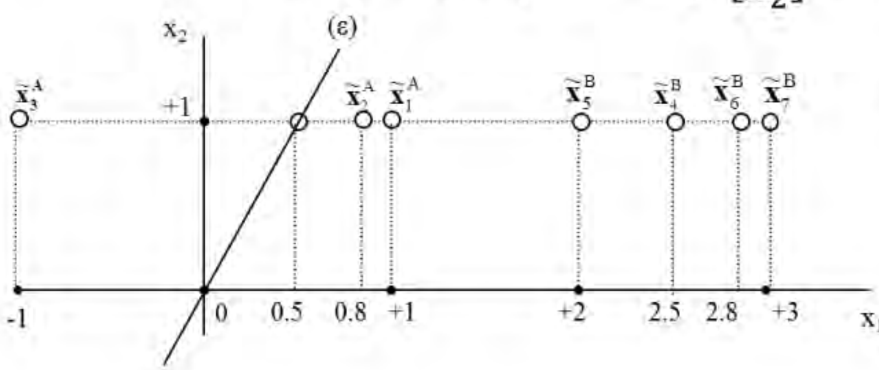
$$w = x^A{}_3 - x^B{}_5 = -3$$

$$w_2 = -\frac{1}{2}\left(x^A{}_3{}^T \cdot x^A{}_3 - x^B{}_5{}^T \cdot x^B{}_5\right) = \frac{3}{2}$$

$$\text{και } w \cdot x_m + w_2 = 0 \Rightarrow x_m = \frac{1}{2}$$

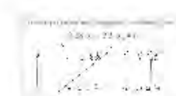
Στο χώρο των επαυξημένων διανυσμάτων:

$$\boldsymbol{W}^T \boldsymbol{X} = 0 \Rightarrow [w_1 \quad w_2] \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow -3x_1 + \frac{3}{2}x_2 = 0 \Rightarrow (\varepsilon) \rightarrow x_2 - 2x_1 = 0$$



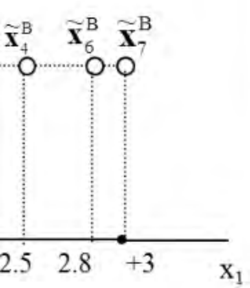
$$D(\tilde{\mathbf{x}}_5^B) = \left[-3, \frac{3}{2}\right]\begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-3) \cdot 2 + \left[\frac{3}{2}\right] \cdot 1 = -4.5 < 0$$

Ακολουθεί έλεγχος κάθε προτύπου με έλεγχο της ποσότητας D(x) για τα επαυξημένα διανύσματα x των προτύπων για κάποια τιμή ρ (έστω εδώ 0.4)



α

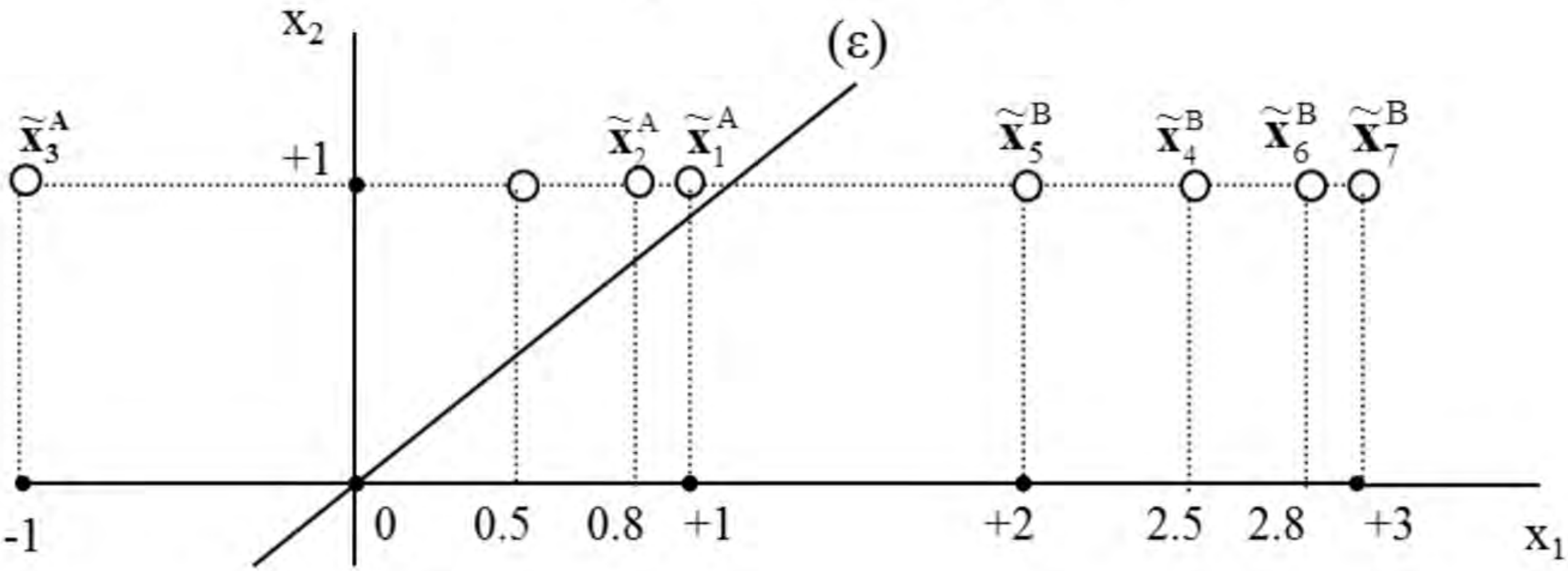
κλάσεις είναι:



t	$\widetilde{\mathbf{x}}=\begin{bmatrix}x_1\\1\end{bmatrix}$	$\mathbf{w}_t=\begin{bmatrix}w_1\\w_2\end{bmatrix}$	$D(\widetilde{\mathbf{x}})=\mathbf{w}^T\cdot\widetilde{\mathbf{x}}$	Πρόσημο κλάσης του $\widetilde{\mathbf{x}}$	Διόρθωση $\mathbf{w}_{t+1}=\mathbf{w}_t\pm\rho\cdot\widetilde{\mathbf{x}}$
0	$\widetilde{\mathbf{x}}_1^A=\begin{bmatrix}1\\1\end{bmatrix}$	$\mathbf{w}_0=\begin{bmatrix}-3\\1.5\end{bmatrix}$	$-3\cdot1+1.5=$ $-1.5<0$	+	$\mathbf{w}_1=\begin{bmatrix}-3\\1.5\end{bmatrix}+0.4\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}-2.6\\1.9\end{bmatrix}$
1	$\widetilde{\mathbf{x}}_4^B=\begin{bmatrix}2.5\\1\end{bmatrix}$	$\mathbf{w}_1=\begin{bmatrix}-2.6\\1.9\end{bmatrix}$	$-2.6\cdot2.5+1.9=$ $-3.5<0$	-	Καμία διόρθωση $\mathbf{w}_2=\mathbf{w}_1$
2	$\widetilde{\mathbf{x}}_2^A=\begin{bmatrix}0.8\\1\end{bmatrix}$	$\mathbf{w}_2=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$-2.6\cdot0.8+1.9=$ $-1.7<0$	+	$\mathbf{w}_3=\begin{bmatrix}-2.6\\1.9\end{bmatrix}+0.4\begin{bmatrix}0.8\\1\end{bmatrix}=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$
3	$\widetilde{\mathbf{x}}_5^B=\begin{bmatrix}2\\1\end{bmatrix}$	$\mathbf{w}_3=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$-2.28\cdot2+2.3<0$	-	Καμία διόρθωση $\mathbf{w}_4=\mathbf{w}_3$
4	$\widetilde{\mathbf{x}}_6^B=\begin{bmatrix}2.8\\1\end{bmatrix}$	$\mathbf{w}_4=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$-2.28\cdot2.8+2.3<0$	-	Καμία διόρθωση $\mathbf{w}_5=\mathbf{w}_4$
5	$\widetilde{\mathbf{x}}_3^A=\begin{bmatrix}-1\\1\end{bmatrix}$	$\mathbf{w}_5=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$(-1.28)\cdot(-1)$ $+2.3>0$	+	Καμία διόρθωση $\mathbf{w}_6=\mathbf{w}_5$
6	$\widetilde{\mathbf{x}}_7^B=\begin{bmatrix}3\\1\end{bmatrix}$	$\mathbf{w}_6=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$-2.28\cdot3+2.3<0$	-	Καμία διόρθωση $\mathbf{w}_7=\mathbf{w}_6$
7	$\widetilde{\mathbf{x}}_2^A=\begin{bmatrix}0.8\\1\end{bmatrix}$	$\mathbf{w}_7=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$(-2.28)\cdot0.8+2.3>$ 0	+	Καμία διόρθωση $\mathbf{w}_8=\mathbf{w}_7$
8	$\widetilde{\mathbf{x}}_4^B=\begin{bmatrix}2.5\\1\end{bmatrix}$	$\mathbf{w}_8=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$(-2.28)\cdot2.5+2.3$ <0	-	Καμία διόρθωση $\mathbf{w}_9=\mathbf{w}_8$
9	$\widetilde{\mathbf{x}}_1^A=\begin{bmatrix}1\\1\end{bmatrix}$	$\mathbf{w}_9=\begin{bmatrix}-2.28\\2.3\end{bmatrix}$	$-2.28\cdot1+2.3>0$	+	Καμία διόρθωση

Τελικά η εξίσωση που διαχωρίζει τις κλάσεις είναι:

$$-2.28 \cdot x_1 + 2.3 \cdot x_2 = 0$$



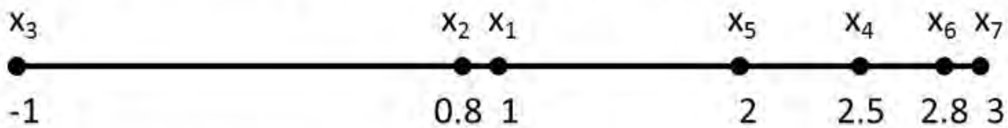
Εκπαίδευση perceptron

Ταξινομητής δύο κλάσεων

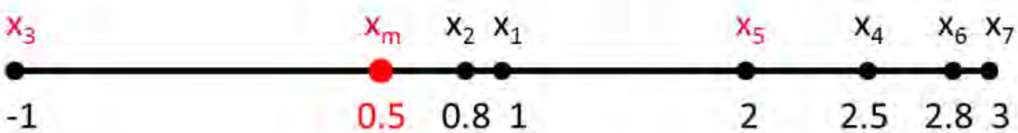
Παράδειγμα σε μία διάσταση

$x^A{}_1 = [1], \quad x^A{}_2 = [0.8], \quad x^A{}_3 = [-1]$

$x^B{}_4 = [2.5], \quad x^B{}_5 = [2], \quad x^B{}_6 = [2.8], \quad x^B{}_7 = [3]$



Επιλέγουμε τυχαία τα $x^A{}_3 = [-1], \quad x^B{}_5 = [2]$



$$x^A{}_3 = [-1], \quad x^B{}_5 = [2] \quad \Rightarrow \; x_m = \frac{1}{2}$$

αφού:

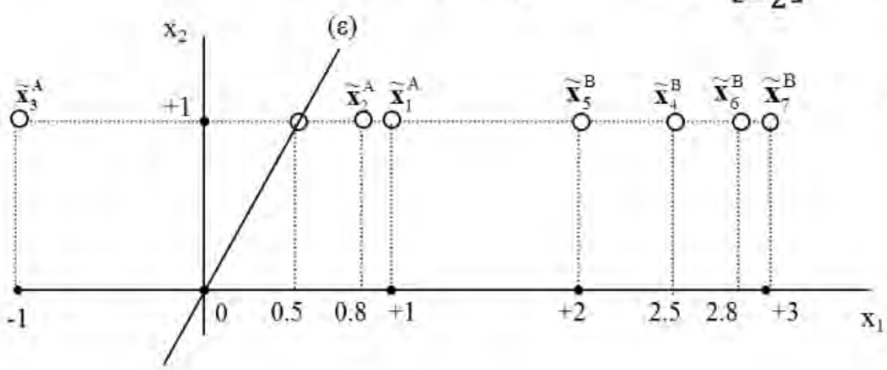
$$w = x^A{}_3 - x^B{}_5 = -3$$

$$w_2 = -\frac{1}{2}\left({x^A{}_3}^T \cdot x^A{}_3 - {x^B{}_5}^T \cdot x^B{}_5\right) = \frac{3}{2}$$

$$\text{και } w \cdot x_m + w_2 = 0 \Rightarrow x_m = \frac{1}{2}$$

Στο χώρο των επαυξημένων διανυσμάτων:

$$\boldsymbol{W}^T \boldsymbol{X} = 0 \Rightarrow [w_1 \quad w_2] \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow -3x_1 + \frac{3}{2}x_2 = 0 \Rightarrow (\varepsilon) \rightarrow x_2 - 2x_1 = 0$$



$$D(\widetilde{\mathbf{x}}_5^B) = \left[\, -3 \,, \frac{3}{2} \right] \begin{bmatrix} 2 \\ 1 \end{bmatrix} = (-3) \cdot 2 + \left[\frac{3}{2}\right] \cdot 1 = -4.5 < 0$$

Ακολουθεί έλεγχος κάθε προτύπου με έλεγχο
της ποσότητας D(x) για τα επαυξημένα διανύσματα
x των προτύπων
για κάποια τιμή ρ (έστω εδώ 0.4)

Εκπαίδευση perceptron

Ταξινομητής δύο κλάσεων

Παράδειγμα σε τρεις διαστάσεις

$$x_1 = \begin{bmatrix}0\\1\\1\end{bmatrix}, x_2 = \begin{bmatrix}0\\0\\1\end{bmatrix}, x_3 = \begin{bmatrix}1\\1\\1\end{bmatrix}, x_4 = \begin{bmatrix}1\\0\\1\end{bmatrix}, \qquad x_1, x_2 \epsilon A, x_3, x_4 \epsilon B$$

Αρχικοποίηση: έστω $w_1 = \begin{bmatrix}0\\0\\0\end{bmatrix}, \qquad w_{N+1} = w_0 = 1$

Θεωρούμε κλάση Α θετικό πρόσημο και Β αρνητικό, άρα:

Αν $x_i \epsilon A, \quad {w_t}^T \cdot x_i \leq 0 \;\Rightarrow\; w_{t+1} = w_t + w_0 \cdot x_i$

Αν $x_i \epsilon B, \quad {w_t}^T \cdot x_i \geq 0 \;\Rightarrow\; w_{t+1} = w_t - w_0 \cdot x_i$

Για λόγους απλότητας υπολογισμών ρ=1

Εποχή 1:

$${w_1}^T \cdot x_1 = [0 \;\; 0 \;\; 0] \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} = 0 \Rightarrow w_2 = w_1 + \rho x_1 = \begin{bmatrix}0\\1\\1\end{bmatrix}$$

$${w_2}^T \cdot x_2 = [0 \;\; 0 \;\; 1] \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 1 \Rightarrow w_3 = w_2$$

$${w_3}^T \cdot x_3 = [0 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 2 \Rightarrow w_4 = w_3 - \rho x_3 = \begin{bmatrix}-1\\0\\0\end{bmatrix}$$

$${w_4}^T \cdot x_4 = [-1 \;\; 0 \;\; 0] \cdot \begin{bmatrix}1\\0\\1\end{bmatrix} = -1 \Rightarrow w_5 = w_4$$

Εποχή 2:

$${w_5}^T \cdot x_1 = [-1 \;\; 0 \;\; 0] \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} = 0 \Rightarrow w_6 = w_5 + \rho x_1 = \begin{bmatrix}-1\\1\\1\end{bmatrix}$$

$${w_6}^T \cdot x_2 = [-1 \;\; 1 \;\; 1] \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 1 \Rightarrow w_7 = w_6$$

$${w_7}^T \cdot x_3 = [-1 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 1 \Rightarrow w_8 = w_7 - \rho x_3 = \begin{bmatrix}-2\\0\\0\end{bmatrix}$$

$${w_8}^T \cdot x_4 = [-2 \;\; 0 \;\; 0] \cdot \begin{bmatrix}1\\0\\1\end{bmatrix} = -2 \Rightarrow w_9 = w_8$$

Εποχή 3:

$${w_9}^T \cdot x_1 = [-2 \;\; 0 \;\; 0] \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} = 0 \Rightarrow w_{10} = w_9 + \rho x_1 = \begin{bmatrix}-2\\1\\1\end{bmatrix}$$

$${w_{10}}^T \cdot x_2 = [-2 \;\; 1 \;\; 1] \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 1 \Rightarrow w_{11} = w_{10}$$

$${w_{11}}^T \cdot x_3 = [-2 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = 0 \Rightarrow w_{12} = w_{11} - \rho x_3 = \begin{bmatrix}-3\\0\\0\end{bmatrix}$$

$${w_{12}}^T \cdot x_4 = [-1 \;\; 0 \;\; 0] \cdot \begin{bmatrix}1\\0\\1\end{bmatrix} = -3 \Rightarrow w_{13} = w_{12}$$

Εποχή 4:

$${w_{13}}^T \cdot x_1 = [-3 \;\; 0 \;\; 0] \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} = 0 \Rightarrow w_{14} = w_{12} + \rho x_1 = \begin{bmatrix}-3\\1\\1\end{bmatrix}$$

$${w_{14}}^T \cdot x_2 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 1 \Rightarrow w_{15} = w_{14}$$

$${w_{15}}^T \cdot x_3 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = -1 \Rightarrow w_{16} = w_{15}$$

$${w_{16}}^T \cdot x_4 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\0\\1\end{bmatrix} = -2 \Rightarrow w_{17} = w_{16}$$

Εποχή 5:

$${w_{17}}^T \cdot x_1 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}0\\1\\1\end{bmatrix} = 2 \Rightarrow w_{18} = w_{17}$$

$${w_{18}}^T \cdot x_2 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}0\\0\\1\end{bmatrix} = 1 \Rightarrow w_{19} = w_{18}$$

$${w_{19}}^T \cdot x_3 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\1\\1\end{bmatrix} = -1 \Rightarrow w_{20} = w_{19}$$

$${w_{20}}^T \cdot x_4 = [-3 \;\; 1 \;\; 1] \cdot \begin{bmatrix}1\\0\\1\end{bmatrix} = -2 \Rightarrow w_{21} = w_{20}$$

Ταξινόμηση σε περισσότερες (των 2) κλάσεις

Περισσότεροι του ενός γραμμικοί ταξινομητές

Προϋπόθεση: γραμμικώς διαχωρίσιμες κλάσεις

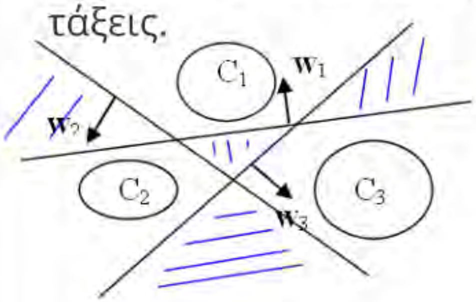
Πρώτη προσέγγιση (M>2 κλάσεις)

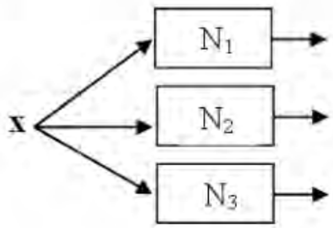
- γραμμικός διαχωρισμός κάθε τάξης από τις υπόλοιπες με M γραμμικές διακριτικές συναρτήσεις

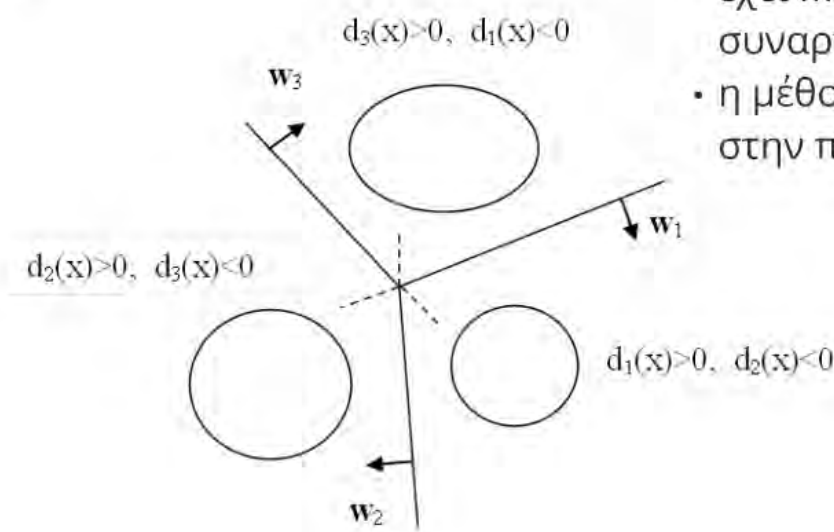
$$d_{\mu}(\boldsymbol{x}) = \boldsymbol{w}_{\mu}^{\mathrm{T}} \cdot \boldsymbol{x} + c \; > 0, \qquad \mu = 1 \ldots \mathrm{M}, \qquad \gamma\iota\alpha\ \boldsymbol{x} \epsilon C_{\mu} \;\Rightarrow\; \boldsymbol{x_i} \epsilon C_{\mu}\ \alpha\nu\ d_{\mu}(\boldsymbol{x_i}) > 0$$

Σε υλοποίηση με ΝΔ και βηματική έξοδο -1/1

- ένα άγνωστο πρότυπο ταξινομείται στην κλάση που ο αντίστοιχος νευρώνας έχει έξοδο 1
- όμως μια τέτοια προσέγγιση είναι δυνατόν να αποδώσει ένα πρότυπο σε καμία ή σε πολλές τάξεις.

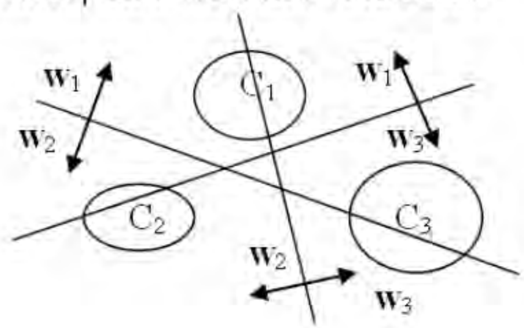






Άλλη προσέγγιση

- διαχωρισμός των κλάσεων ανά δύο αδιαφορώντας για τις υπόλοιπες.
- ένα πρότυπο αποδίδεται σε εκείνη την κλάση που έχει M-1 θετικές τιμές στις διακριτικές συναρτήσεις που την χωρίζουν από τις υπόλοιπες
- η μέθοδος αυτή είναι υπολογιστικά πολύπλοκη στην περίπτωση των πολλών κλάσεων



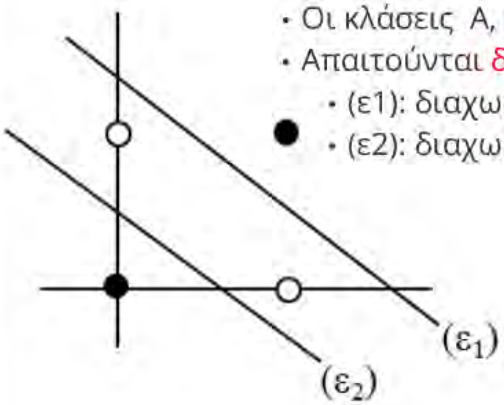
Πολυεπίπεδα perceptron

Ταξινομητές πολλών επιπέδων

a	b	a XOR b
0	0	**0**
0	1	**1**
1	0	**1**
1	1	**0**

Πρότυπα: $\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right\}$

Η πράξη XOR ορίζει τις κλάσεις $A=\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1\\1\end{bmatrix}\right\}, B=\left\{\begin{bmatrix}0\\1\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix}\right\}$

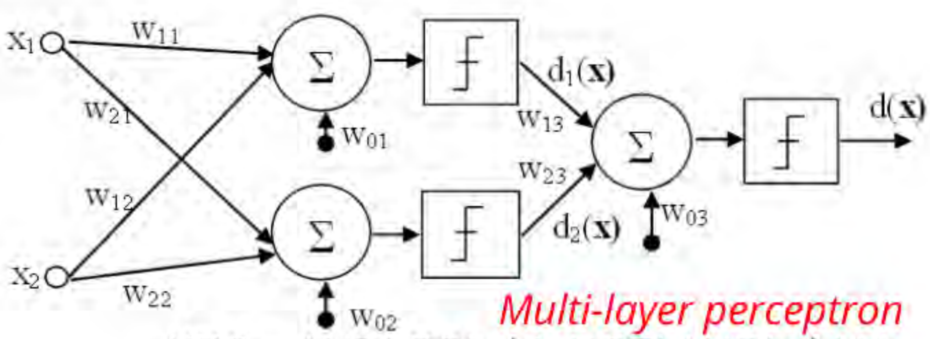


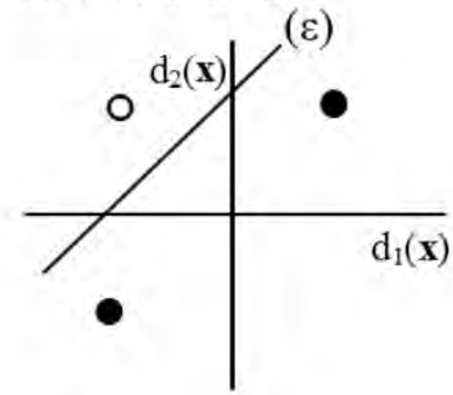
- Οι κλάσεις A, B δε διαχωρίζονται με μια ευθεία
- Απαιτούνται δύο ευθείες (ε1), (ε2) ως εξής:
  - (ε1): διαχωρίζει το [1,1] από τα υπόλοιπα (ταξινομητής T1)
  - (ε2): διαχωρίζει το [0,0] από τα υπόλοιπα (T2)

a	b	T1→d_1(x)	T2→d_2(x)	ΚΛΑΣΗ
0	0	-1	-1	A
0	1	-1	+1	B
1	0	-1	+1	B
1	1	+1	+1	A

- Οι τιμές των d1(x), d2(x) προσδιορίζουν τις κλάσεις A και B αποτελώντας ένα νέο σύνολο προτύπων Φ = {[-1,-1], [-1,+1], [+1,+1]}
- η περίπτωση [+1,-1] είναι αδύνατη και δεν περιλαμβάνεται στο χώρο Φ
- Οι κλάσεις πλέον προσδιορίζονται από το σημείο [-1,+1] του χώρου Φ που διαχωρίζεται γραμμικά με την ευθεία (ε) από τα διανύσματα [-1,-1],[+1,+1]

Ένας γραμμικός ταξινομητής T μπορεί να προσδιορίσει την (ε).
Σ' αυτόν:
αν $d_T(\boldsymbol{x}) > 0 \Rightarrow \boldsymbol{x}\in A$
αν $d_T(\boldsymbol{x}) < 0 \Rightarrow \boldsymbol{x}\in B$





Multi-layer perceptron

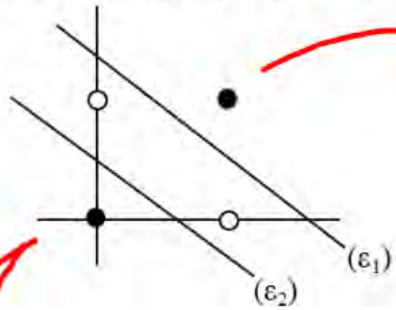
XOR = AND(OR(a,b), NOT(AND(a,b)))

Πολυεπίπεδα perceptron

Ταξινομητές πολλών επιπέδων

Να σχεδιαστεί ταξινομητής για την εκμάθηση της λογικής πράξης XOR

a	b	a XOR b
0	0	**0**
0	1	**1**
1	0	**1**
1	1	**0**



αν η (ε_1) τέμνει τους άξονες στα σημεία $\begin{bmatrix} 5/4 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 5/4 \end{bmatrix}$

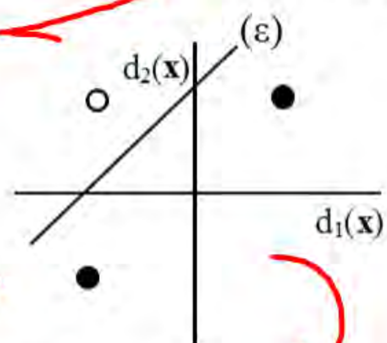
τότε η εξίσωσή της είναι: $(\varepsilon_1) \rightarrow \frac{4}{5}x_1 + \frac{4}{5}x_2 - 1 = 0$

ή $(\varepsilon_1) \rightarrow 0.8x_1 + 0.8x_2 - 1 = 0$

αν η (ε_2) τέμνει τους άξονες στα σημεία $\begin{bmatrix} 1/2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}$

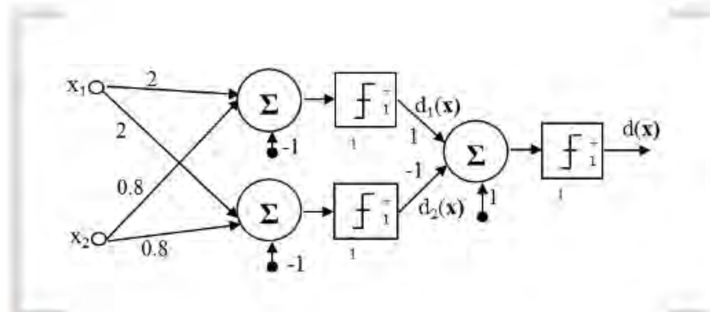
τότε η εξίσωσή της είναι: $(\varepsilon_2) \rightarrow 2x_1 + 2x_2 - 1 = 0$

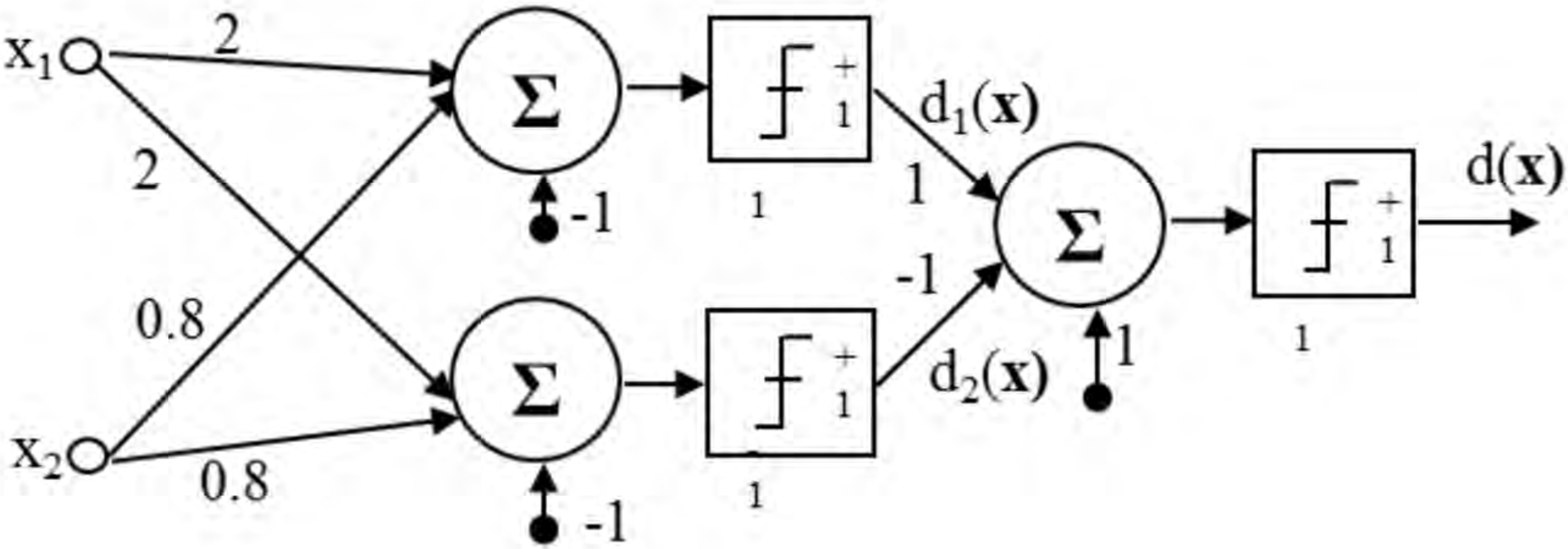
a	b	T1→$d_1(x)$	T2→$d_2(x)$	ΚΛΑΣΗ
0	0	-1	-1	A
0	1	-1	+1	B
1	0	-1	+1	B
1	1	+1	+1	A



κατάλληλη ευθεία (ε) τέμνει τους άξονες στα σημεία $\begin{bmatrix} -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

τότε η εξίσωσή της είναι: $(\varepsilon) \rightarrow -x_1 + x_2 - 1 = 0 \Rightarrow (\varepsilon) \rightarrow x_1 - x_2 + 1 = 0$





x1
x2
2
2
0.8
0.8
Σ
Σ
-1
-1
+
1
1
+
1
1
d1(x)
d2(x)
1
-1
Σ
1
+
1
1
d(x)