

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ
Π.Σ. ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ Τ.Ε.

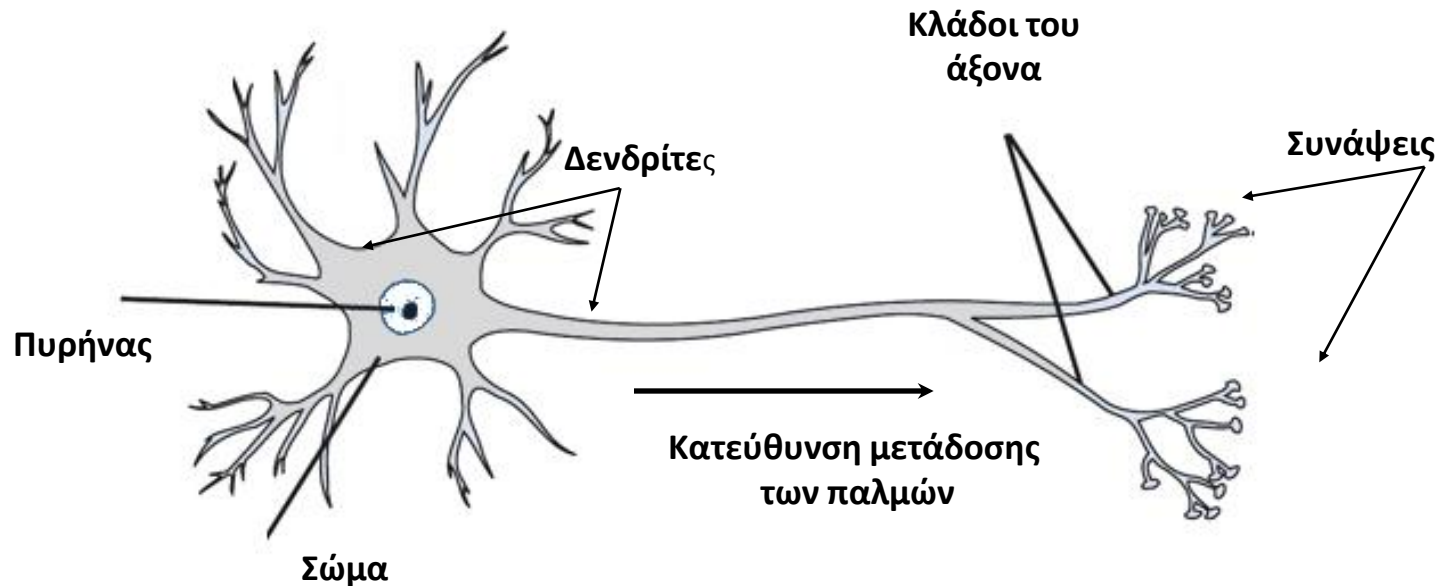
ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ
ΜΑΘΗΜΑ 3^ο

ΣΤΑΥΡΟΣ ΑΔΑΜ

ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ 2020-2021

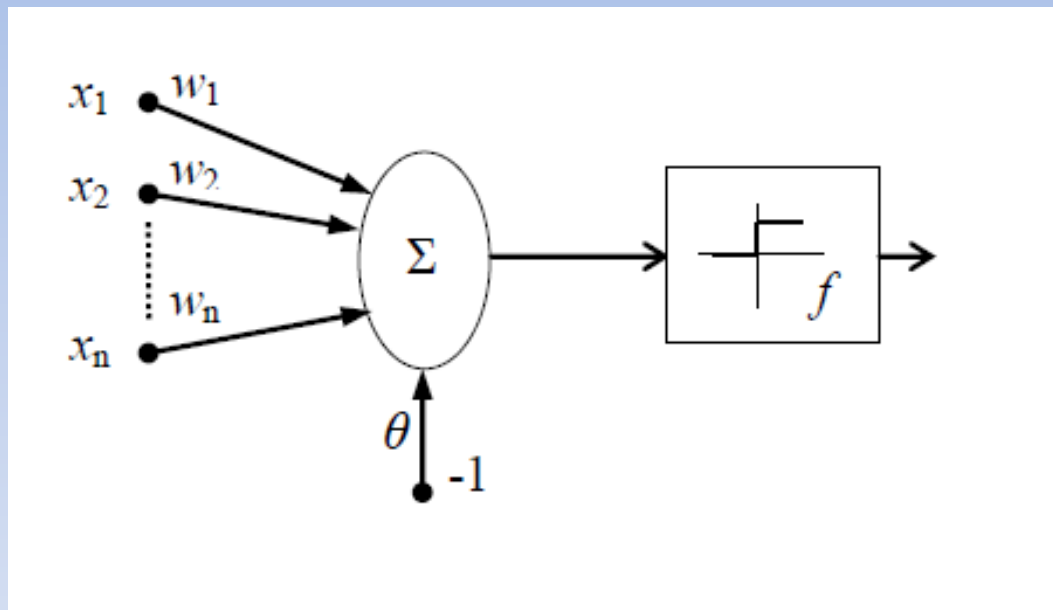
ΒΙΟΛΟΓΙΚΟΣ ΚΑΙ ΤΕΧΝΗΤΟΣ ΝΕΥΡΩΝΑΣ

Τυπική δομή βιολογικού νευρώνα



ΒΙΟΛΟΓΙΚΟΣ ΚΑΙ ΤΕΧΝΗΤΟΣ ΝΕΥΡΩΝΑΣ

Ο νευρώνας των McCulloch – Pitts



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

$$y = f(u - \theta)$$

$$s = u - \theta$$

$$f(s) = \begin{cases} 0, & \alpha \nu s \leq 0 \\ 1, & \alpha \nu s > 0 \end{cases}$$

Ο νευρώνας των McCulloch – Pitts

Παραδοχές

- Οι είσοδοι αναπαρίστανται ως ένα διάνυσμα τιμών που ονομάζεται **επαυξημένο διάνυσμα**:

$$\mathbf{x} = [x_0, x_1, \dots, x_n]^T$$

- Η επί πλέον τιμή x_0 αντιστοιχεί στο σήμα εισόδου του bias και συνήθως τίθεται ίσο με 1
- Στο διάνυσμα αυτό αντιστοιχεί **το επαυξημένο διάνυσμα** των βαρών:

$$\mathbf{w} = [w_0, w_1, \dots, w_n]^T$$

Παραδοχές (συνέχεια)

- Κατά συνέπεια το άθροισμα:

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

γράφεται:

$$u = \sum_{i=1}^n w_i x_i + w_0 = \sum_{i=0}^n w_i x_i$$

και το u αντικαθιστά τη μεταβλητή s

Σχετικά με τη Μάθηση

- Η μάθηση από παραδείγματα είναι αποτέλεσμα της διαδικασίας εκπαίδευσης. Στόχος της είναι ο καθορισμός των τιμών των συναπτικών βαρών ώστε το δίκτυο να επιτελεί μια νοήμονα λειτουργία η οποία περιγράφεται άτυπα από τα παραδείγματα.
- Ικανότητα γενίκευσης. Η μάθηση είναι επιτυχής όταν το ΤΝΔ είναι ικανό να επιτελεί την επιθυμητή λειτουργία όχι μόνο στα γνωστά παραδείγματα αλλά και σε άγνωστα, επιδεικνύοντας έτσι ικανότητα γενίκευσης.

Σχετικά με τη Μάθηση

- Όταν τα παραδείγματα εκπαίδευσης είναι αντιπροσωπευτικά της λειτουργίας για εκμάθηση τότε τα ΤΝΔ έχουν αποδείξει ότι αποτελούν συστήματα με αξιόλογη γενικευτική ικανότητα.

Μάθηση με παραδείγματα

Διάφορες προσεγγίσεις

- Μάθηση με επίβλεψη (supervised learning)
- Μάθηση χωρίς επίβλεψη (unsupervised learning)
- Μάθηση με ενίσχυση (reinforcement learning)
- Μάθηση με ημι-επίβλεψη (semi-supervised learning)

Μάθηση με επίβλεψη (supervised learning)

- Τα παραδείγματα του προβλήματος εμφανίζονται με τη μορφή συνόλου από ζεύγη της μορφής (είσοδος, επιθυμητή έξοδος ή στόχος)

$$\{(x^p, t^p), p = 1 \dots P\}$$

- Κάθε ζεύγος μπορεί να θεωρηθεί ως (ερώτηση, ορθή απάντηση)

Σημειώστε ότι $\forall p, \mathbf{x}^p = [x_1^p, x_2^p, \dots, x_n^p]^T$

Μάθηση με επίβλεψη (supervised learning)

- Αυτό που κατ'αρχήν μαθαίνει το ΤΝΔ είναι να αντιστοιχίζει σωστά σε κάθε είσοδο την αντίστοιχη έξοδο. Ο βαθμός στον οποίο αυτή η αντιστοίχιση είναι σωστή χαρακτηρίζει και την επιτυχία της μάθησης
- Ο όρος μάθηση με επίβλεψη παραπέμπει στο ανθρώπινο τρόπο μάθησης όπου ένας επιβλέπων «ο δάσκαλος» εποπτεύει τη διαδικασία μάθησης θέτοντας ερωτήματα και δίνοντας τις σωστές απαντήσεις

PERCEPTRON

Βασικά στοιχεία

- Ένα perceptron είναι ένα ΤΝΔ που αποτελείται από νευρώνες McCulloch-Pitts
- Περιλαμβάνει ένα ή περισσότερους νευρώνες
- Όλοι οι νευρώνες δέχονται στην είσοδο τις τιμές που περιγράφουν τα πρότυπα $x^p, p = 1 \dots P$
- Επί πλέον στο perceptron παρουσιάζεται και η τιμή της κατηγορίας t^p που αντιστοιχεί στο πρότυπο x^p
- Με χρήση ενός αλγορίθμου που ρυθμίζει τα συναπτικά βάρη το perceptron μαθαίνει να αντιστοιχεί σε κάθε πρότυπο την κατηγορία στην οποία ανήκει

PERCEPTRON

- Το διάνυσμα των δεδομένων δίνεται ή μετατρέπεται σε επαυξημένη μορφή:

$$\forall p, \mathbf{x}^p = [x_0^p, x_1^p, x_2^p, \dots, x_n^p]^T$$

- Η τιμή κάθε απάντησης (στόχου) είναι:

a) $\{-1, 1\}$

ενώ πολλές φορές χρησιμοποιείται και

b) $\{0, 1\}$

Μάθηση με επίβλεψη (supervised learning)

- Στην πρώτη περίπτωση η συνάρτηση ενεργοποίησης του τεχνητού νευρώνα είναι,

η συνάρτηση:
$$f(u) = \begin{cases} -1, & \text{αν } u \leq 0 \\ 1, & \text{αν } u > 0 \end{cases}$$

ή ακόμη:
$$f(u) = \begin{cases} -1, & \text{αν } u < 0 \\ 1, & \text{αν } u \geq 0 \end{cases}$$

Ενώ στη δεύτερη περίπτωση είναι η συνάρτηση:

$$f(u) = \begin{cases} 0, & \text{αν } u \leq 0 \\ 1, & \text{αν } u > 0 \end{cases}$$

PERCEPTRON

Τρόπος λειτουργίας

1. Για κάθε πρότυπο, $p = 1 \dots P$

υπολογίζεται η συνολική είσοδος, δηλ. το επίπεδο ενεργοποίησης του νευρώνα:

$$u^p = \sum_{i=1}^n w_i x_i^p + w_0 = \sum_{i=0}^n w_i x_i^p$$

2. Υπολογισμός της εξόδου του νευρώνα με χρήση της συνάρτησης ενεργοποίησης (activation function)

$$y^p = f(u^p)$$

PERCEPTRON

Τρόπος λειτουργίας (συνέχεια)

3. Το εσωτερικό γινόμενο $u^p = \mathbf{w}^T \mathbf{x}^p$

υπολογίζεται με τα επαυξημένα διανύσματα για κάθε πρότυπο

$$\mathbf{x}^p = [x_0^p, x_1^p, \dots, x_n^p]^T$$

$$\mathbf{w} = [w_0, w_1, \dots, w_n]^T$$

Εκπαίδευση του perceptron

Δεδομένα:

- Το σύνολο των παραδειγμάτων εκπαίδευσης

$$X = \{(\mathbf{x}^p, C(\mathbf{x}^p)), p = 1..P\}$$

- Εδώ, κάθε παράδειγμα αποτελείται από:
 - ένα επαυξημένο διάνυσμα $n+1$ διαστάσεων:
για το p πρότυπο
 - το διακριτικό της κατηγορίας : $C(\mathbf{x}^p) \equiv C_1$ ή C_2
του προτύπου

PERCEPTRON

Εκπαίδευση του perceptron

Δεδομένα:

- Το διακριτικό της κατηγορίας αποτελεί συμβολική πληροφορία και θα πρέπει να μετατραπεί σε αριθμητική δηλ. -1 και 1
- Άρα, απαιτείται ο καθορισμός των τιμών στόχων για κάθε κατηγορία και κατά συνέπεια για κάθε πρότυπο:

$$C_1 \rightarrow t = 1 \text{ και } C_2 \rightarrow t = -1$$

- Τα δεδομένα είναι πλέον:

$$X = \{(\mathbf{x}^p, t^p), p = 1..P\}$$

PERCEPTRON

Εκπαίδευση του perceptron

Ανακεφαλαιώνοντας:

- Παραδείγματα εκπαίδευσης: $X = \{(\mathbf{x}^p, t^p), p = 1..P\}$ που ανήκουν σε δύο κατηγορίες
- Κάθε κατηγορία σηματοδοτείται από μία τιμή στόχο
- Η εκπαίδευση συνίσταται στον καθορισμό εκείνων των τιμών για τα βάρη του perceptron ώστε για κάθε πρότυπο p η τιμή της εξόδου του perceptron να είναι ίση με την τιμή του στόχου. Δηλαδή να ισχύει

$$y^p = t^p$$

Εκπαίδευση του perceptron

Ανακεφαλαιώνοντας:

- Αρχικά επιλέγεται ένα τυχαίο διάνυσμα βαρών που δίνει μια τυχαία θέση του συνόρου απόφασης δηλ. της διαχωριστικής ευθείας (επιπέδου ή υπερεπιπέδου)
- Για την τυχαία αυτή θέση του συνόρου απόφασης κάποια από τα πρότυπα ενδέχεται να ταξινομούνται σωστά και κάποια λάθος. Η θέση του συνόρου απόφασης διορθώνεται με αλλαγή των συναπτικών βαρών και υπολογίζεται η νέα θέση του

PERCEPTRON

Εκπαίδευση του perceptron

- Τα πρότυπα ή αλλιώς παραδείγματα παρουσιάζονται ένα προς ένα στο perceptron με κυκλική σειρά:

Επανάληψη	1	2	...	P	P+1	P+2	...	2P	2P+1	...
Πρότυπο	1	2	...	P	1	2	...	P	1	...
	Εποχή 1				Εποχή 2				Εποχή 3	

PERCEPTRON

Εκπαίδευση του perceptron

Για κάθε πρότυπο $\mathbf{x}^p$

- Σωστή ταξινόμηση του προτύπου σημαίνει ότι το πρότυπο βρίσκεται στη σωστή πλευρά σε σχέση με το σύνορο απόφασης (ευθεία, επίπεδο ή υπερεπίπεδο).
→ **Άρα καμμία αλλαγή στα βάρη του perceptron**
- Λανθασμένη ταξινόμηση του προτύπου σημαίνει ότι το πρότυπο βρίσκεται στη λάθος πλευρά σε σχέση με το υπερεπίπεδο. Άρα απαιτείται να αλλάξει θέση το υπερεπίπεδο.
→ **Άρα μεταβολή των τιμών των βαρών**

Εκπαίδευση του perceptron

Σχετικά με τη μεταβολή των βαρών:

- Η μεταβολή πρέπει να μετακινήσει το σύνορο απόφασης προς την κατεύθυνση διόρθωσης του σφάλματος ώστε,
 - είτε το λάθος ταξινομημένο πρότυπο να βρεθεί στη σωστή πλευρά του συνόρου απόφασης,
 - είτε να μειωθεί η απόσταση που χωρίζει το πρότυπο από την σωστή πλευρά του συνόρου απόφασης.

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Συμβολισμοί:

- k η επανάληψη του αλγορίθμου
- $\mathbf{w}(k)$ το επαυξημένο διάνυσμα βαρών στην k επανάληψη
- $\mathbf{x}^p(k)$ το επαυξημένο διάνυσμα που αντιστοιχεί στο πρότυπο p στην k επανάληψη και ο αντίστοιχος στόχος $t^p(k)$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Αρχικοποίηση:

1. Θέτουμε $k = 0$
2. Επιλέγουμε τυχαία τιμές για το διάνυσμα βαρών σε ένα διάστημα όπως π.χ. το διάστημα $[-1,1]$
 $w(0)$
3. Δίνουμε τιμή στο συντελεστή μάθησης η , ($\eta > 0$)
4. $k = k + 1$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Σε κάθε επανάληψη $k > 0$:

1. Επιλέγεται ένα πρότυπο και δίνεται στην είσοδο του perceptron
2. Υπολογίζονται η διέγερση $u^p(k)$ και η έξοδος $y^p(k)$
3. Αν $y^p(k) \neq t^p(k)$ τότε τα συναπτικά βάρη τροποποιούνται με χρήση της διανυσματικής σχέσης:

$$\mathbf{w}(k) = \mathbf{w}(k-1) - \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)$$

4. Αλλιώς τα βάρη μένουν αμετάβλητα

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Τερματισμός του αλγορίθμου:

1. Αν για μία εποχή δεν υπάρξει αλλαγή στα βάρη τότε ο αλγόριθμος τερματίζει

Αλγόριθμος Εκπαίδευσης του perceptron

Σύγκλιση:

- Από το Θεώρημα σύγκλισης του κανόνα perceptron αποδεικνύεται ότι αν οι δύο κατηγορίες είναι γραμμικά διαχωρίσιμες τότε ο αλγόριθμος συγκλίνει σε πεπερασμένο αριθμό επαναλήψεων έχοντας προσδιορίσει τιμές για τα βάρη που ορίζουν ένα υπερεπίπεδο που αποτελεί το διαχωριστικό σύνορο μεταξύ των δύο κατηγοριών
- Αν οι κατηγορίες δεν είναι γραμμικά διαχωρίσιμες τότε ο αλγόριθμος αυτός «ταλαντώνεται» ανάμεσα στις δύο κατηγορίες

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών:

- Έστω ότι στην k επανάληψη ισχύει ότι: $t^p(k) = 1$
και $y^p(k) = -1$
- Κατά συνέπεια $(t^p(k) - y^p(k)) = 2$
$$y^p(k) = -1 \Rightarrow u^p(k) < 0$$
- Άρα χρειάζεται $u^p(k+1) > u^p(k)$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών (συνέχεια)

- Μετά την ενημέρωση των βαρών για το $\mathbf{x}^p(k)$ πρότυπο θα ισχύει ότι:

$$u^p(k+1) = \mathbf{w}(k+1)^T \mathbf{x}^p(k) = \\ \mathbf{w}(k)^T \mathbf{x}^p(k) + \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)^T \mathbf{x}^p(k)$$

Άρα

$$u^p(k+1) = u^p(k) + \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)^T \mathbf{x}^p(k)$$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών (συνέχεια)

- Δεδομένου ότι:

$$\mathbf{x}^p(k)^T \mathbf{x}^p(k) = \|\mathbf{x}^p(k)\|^2 > 0, \eta > 0 \text{ και}$$

$$(t^p(k) - y^p(k)) > 0$$

- Συμπεραίνουμε πως,

$$u^p(k+1) > u^p(k)$$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών:

- Αν στην k επανάληψη ισχύει ότι:

$$t^p(k) = -1 \quad \text{και} \quad y^p(k) = 1$$

- Τότε $(t^p(k) - y^p(k)) = -2$

- Επί πλέον $y^p(k) = 1 \Rightarrow u^p(k) > 0$

- Άρα χρειάζεται $u^p(k+1) < u^p(k)$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών:

- Μετά την ενημέρωση των βαρών για το $\mathbf{x}^p(k)$

θα ισχύει ότι:

$$u^p(k+1) = \mathbf{w}(k+1)^T \mathbf{x}^p(k) = \\ \mathbf{w}(k)^T \mathbf{x}^p(k) + \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)^T \mathbf{x}^p(k)$$

και

$$u^p(k+1) = u^p(k) + \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)^T \mathbf{x}^p(k)$$

PERCEPTRON

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών:

- Μετά την ενημέρωση των βαρών για το $\mathbf{x}^p(k)$

θα ισχύει ότι:

$$u^p(k+1) = \mathbf{w}(k+1)^T \mathbf{x}^p(k) = \\ \mathbf{w}(k)^T \mathbf{x}^p(k) + \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)^T \mathbf{x}^p(k)$$

και

$$u^p(k+1) = u^p(k) + \eta(t^p(k) - y^p(k))\mathbf{x}^p(k)^T \mathbf{x}^p(k)$$

Αλγόριθμος Εκπαίδευσης του perceptron

Επεξήγηση της διόρθωσης των βαρών (συνέχεια)

- Δεδομένου ότι:

$$\mathbf{x}^p(k)^T \mathbf{x}^p(k) = \|\mathbf{x}^p(k)\|^2 > 0, \eta > 0 \text{ και}$$

$$(t^p(k) - y^p(k)) < 0$$

- Συμπεραίνουμε ότι: $u^p(k+1) < u^p(k)$

PERCEPTRON

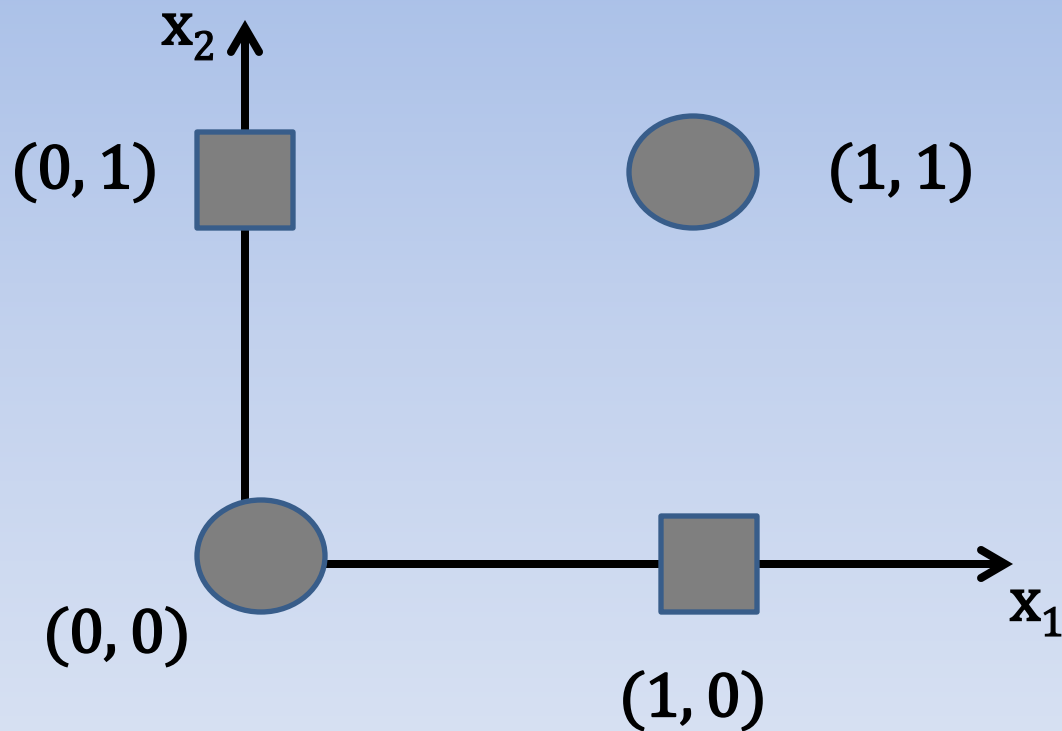
Άλλοι τύποι προβλημάτων

Η λογική συνάρτηση XOR

x_1	x_2	Έξοδος
0	0	0
0	1	1
1	0	1
1	1	0

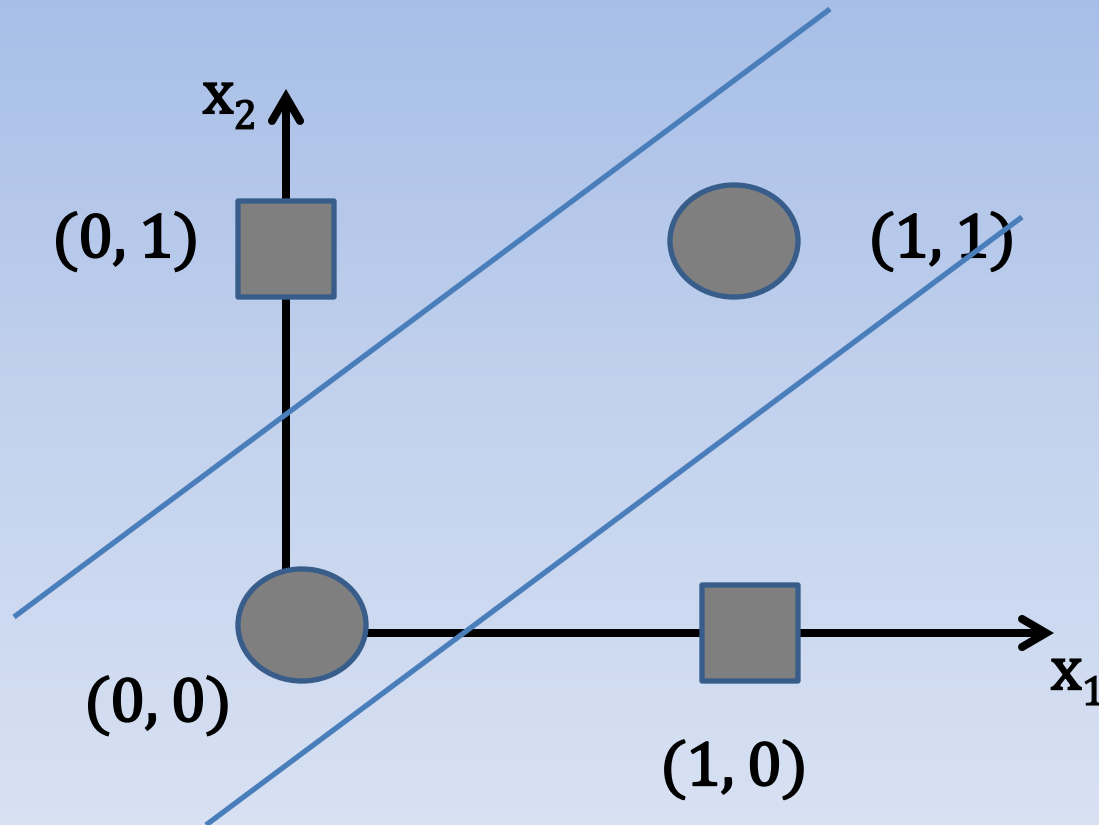
PERCEPTRON

Γραφικά:



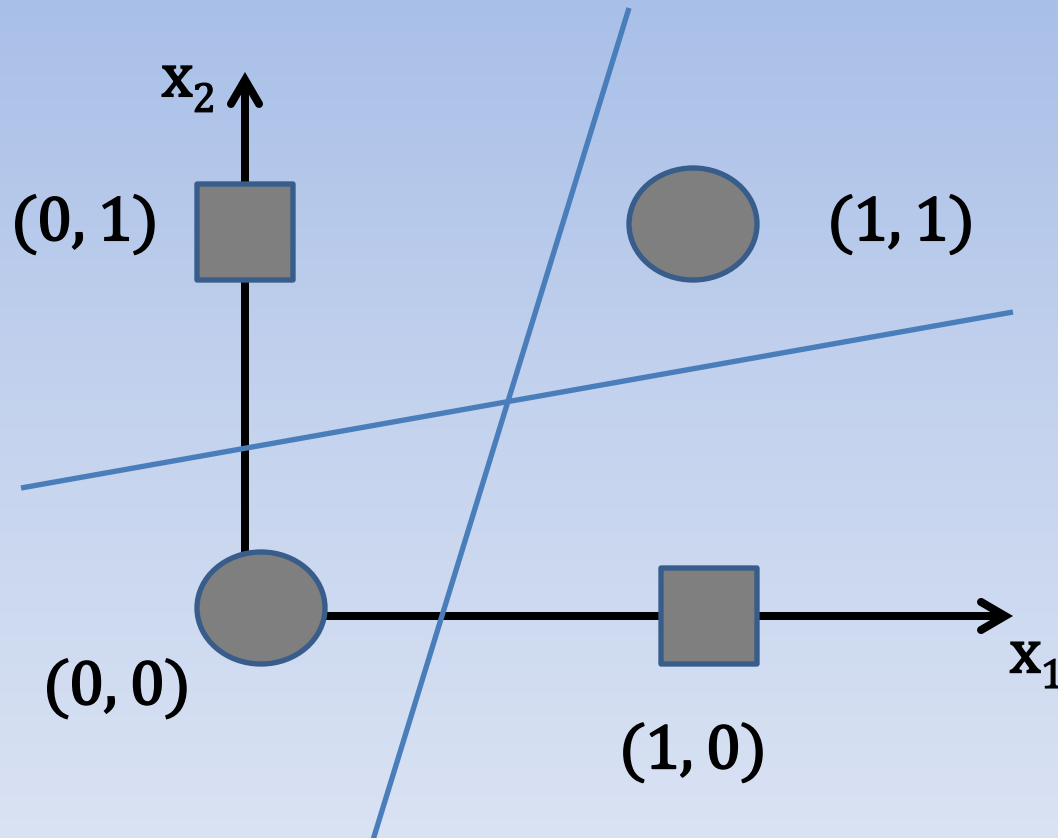
PERCEPTRON

Διαχωρισμός με σύνορο απόφασης δύο ευθείες



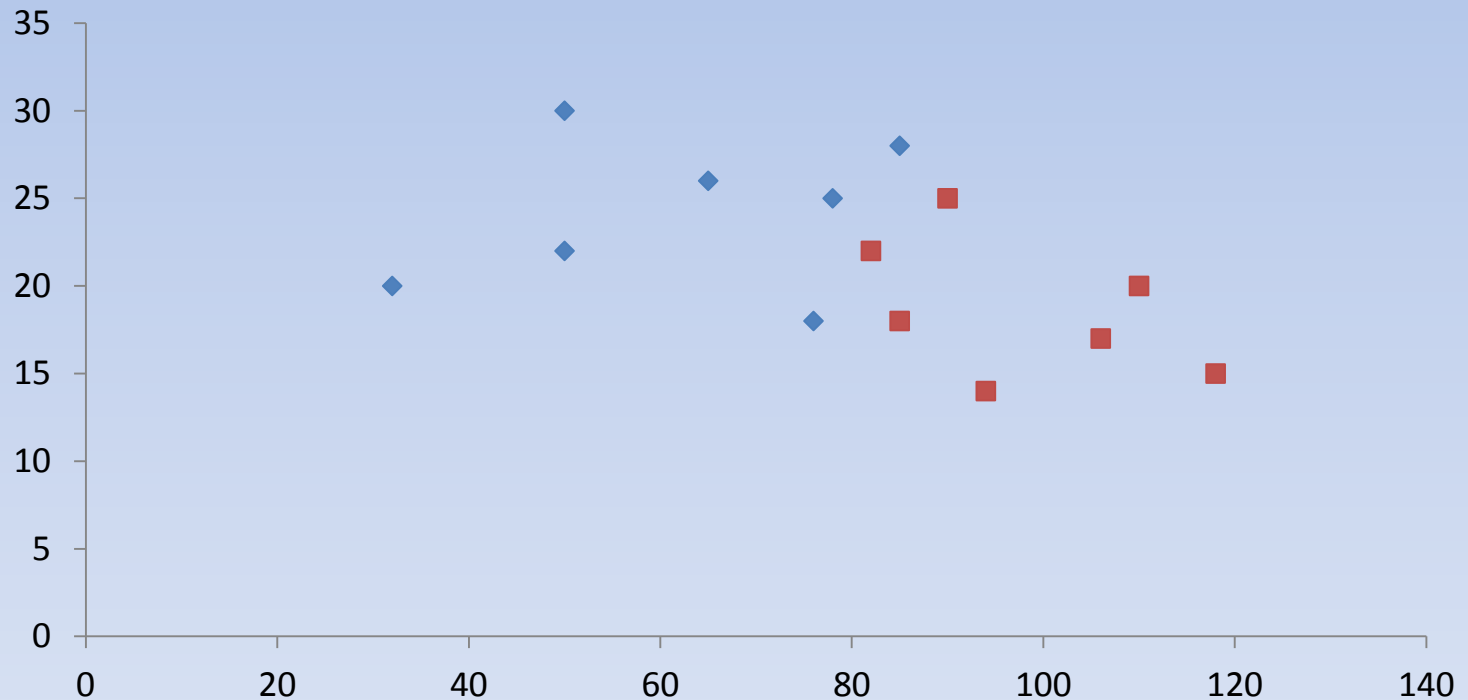
PERCEPTRON

Διαχωρισμός με σύνορο απόφασης δύο ευθείες



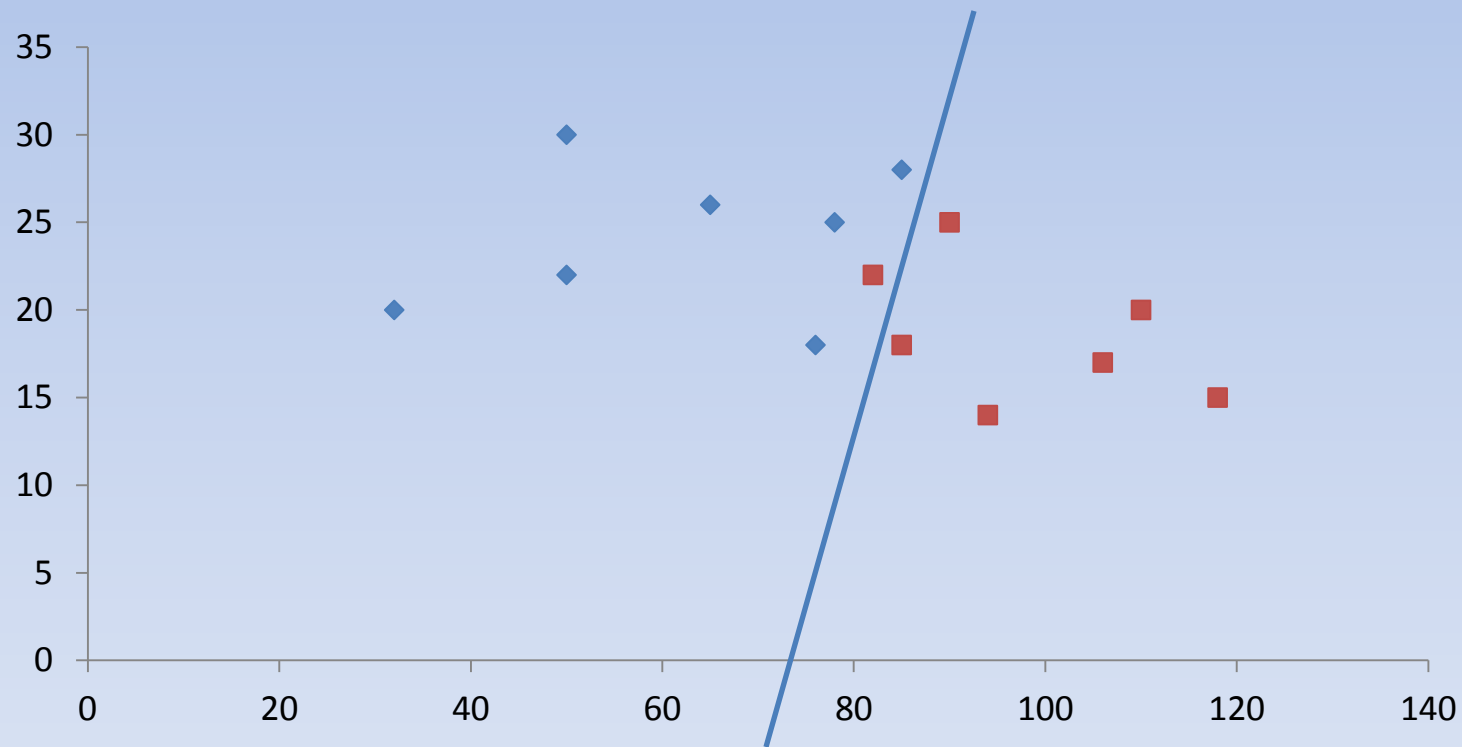
PERCEPTRON

Δύο μη γραμμικά διαχωρίσιμες κατηγορίες, 2-διαστάσεις



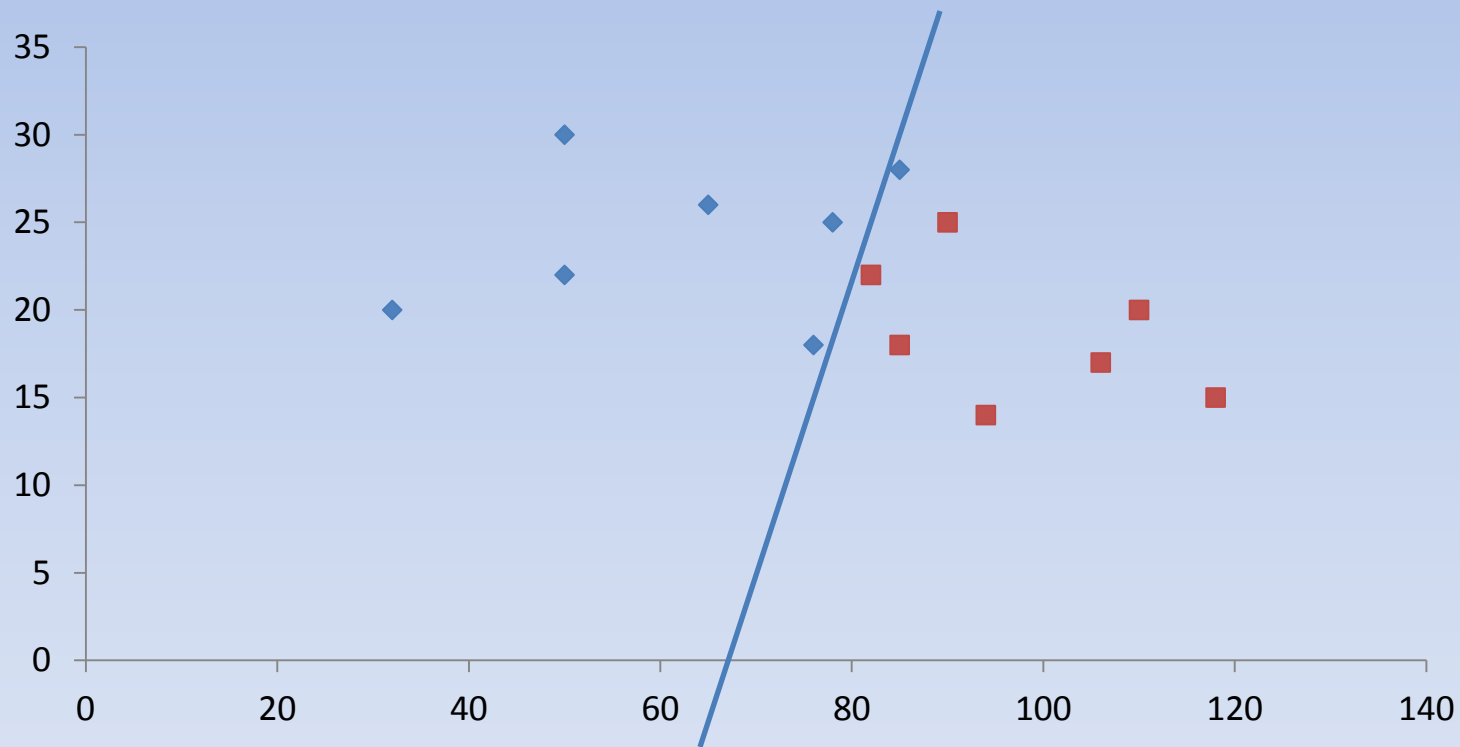
PERCEPTRON

Δύο μη γραμμικά διαχωρίσιμες κατηγορίες, 2-διαστάσεις



PERCEPTRON

Δύο μη γραμμικά διαχωρίσιμες κατηγορίες, 2-διαστάσεις



PERCEPTRON

Δύο μη γραμμικά διαχωρίσιμες κατηγορίες, 2-διαστάσεις

- Υπάρχει η δυνατότητα να εκπαιδεύσουμε ένα perceptron ώστε να μπορεί να διαχωρίσει δύο μη γραμμικά διαχωρίσιμες κατηγορίες με ένα γραμμικό σύνορο απόφασης έτσι ώστε το σφάλμα που θα προκύπτει να είναι αμελητέο για το συγκεκριμένο πρόβλημα, δηλ. να μη μας ενδιαφέρει.
- Ο αλγόριθμος αυτός είναι ο αλγόριθμος pocket perceptron (Gallant, 1990)