Εξελικτικοί Αλγόριθμοι Βελτιστοποίησης

Τεχνητά Νευρωνικά Δίκτυα
(Artificial Neural Networks - ANNs)

Υπεύθυνος μαθήματος : Αθανάσιος Χασιακός
Επιμέλεια παρουσίασης : Παναγιώτης Τσίκας

Τεχνητή Νοημοσύνη - Artificial Intelligence

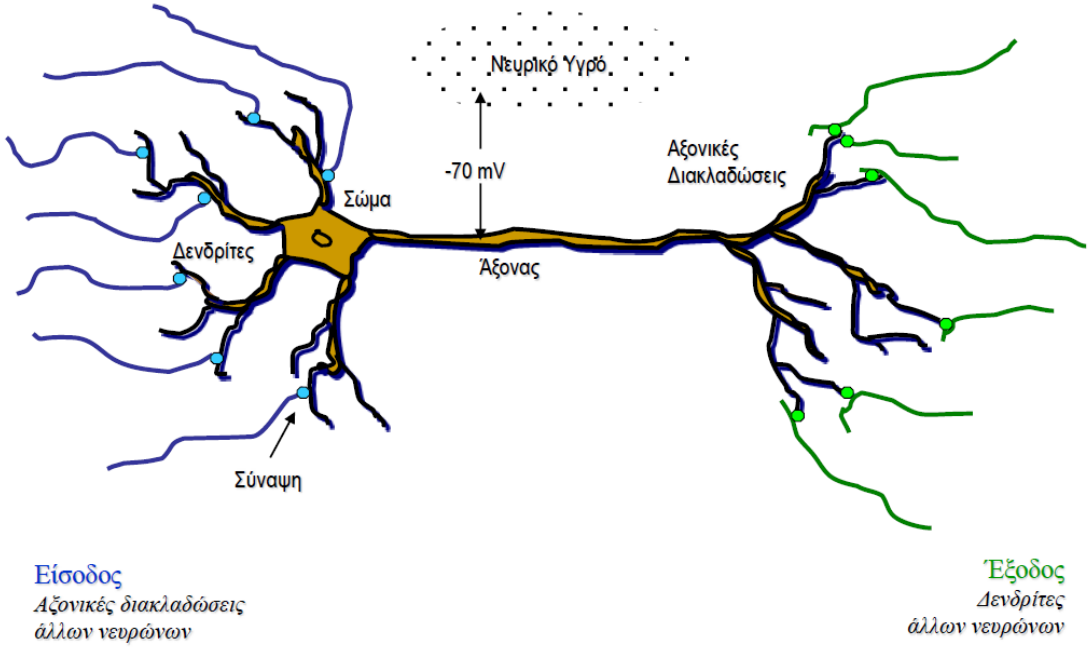
Τεχνητή νοημοσύνη είναι η υλοποίηση προγραμμάτων τα οποία είναι ικανά να μιμηθούν τις ανθρώπινες γνωστικές ικανότητες, όπως η επίλυση προβλημάτων, η αντίληψη μέσω της όρασης, η μάθηση, η εξαγωγή συμπερασμάτων, η κατανόηση φυσικής γλώσσας.

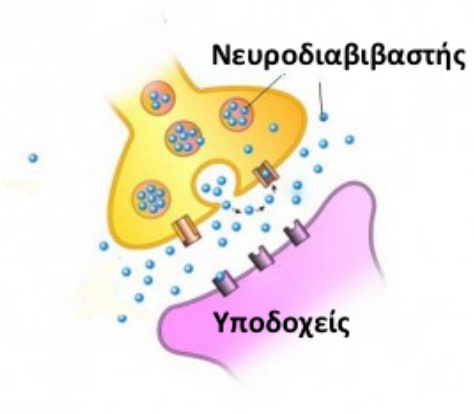
Ένα σύστημα τεχνητής νοημοσύνης πρέπει να είναι ικανό να κάνει τρία πράγματα:
(1) να αποθηκεύει γνώση,
(2) να εφαρμόζει την αποθηκευμένη γνώση,
(3) να αποκτά νέα γνώση μέσω εμπειρίας.

Η τεχνητή νοημοσύνη χωρίζεται σε δυο κατηγορίες
α) Συμβολική
β) Μη συμβολική π.χ. τα Τεχνητά Νευρωνικά Δίκτυα.

Φυσικός Νευρώνας

Η ικανότητα μάθησης και μνήμης που παρουσιάζει ο εγκέφαλος οφείλεται στην ικανότητα των συνάψεων να μεταβάλουν την αγωγιμότητα τους. Τα ηλεκτρικά σήματα που εισέρχονται στο σώμα μέσω των δενδριτών συνδυάζονται και αν το αποτέλεσμα ξεπερνά κάποια τιμή κατωφλίου το σήμα διαδίδεται με τη βοήθεια του άξονα προς άλλους νευρώνες.



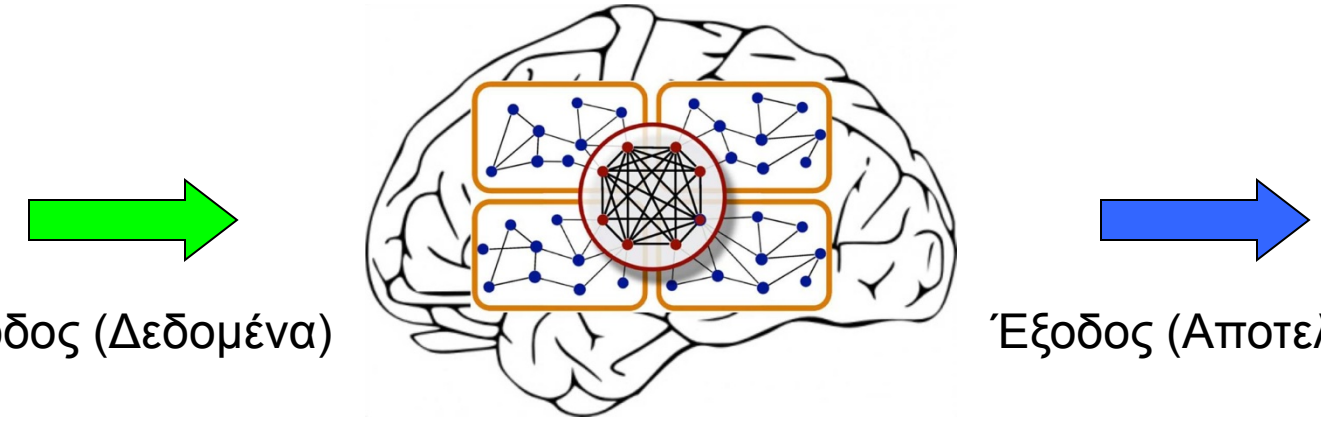


Τεχνητά Νευρωνικά Δίκτυα - Artificial Neural Networks

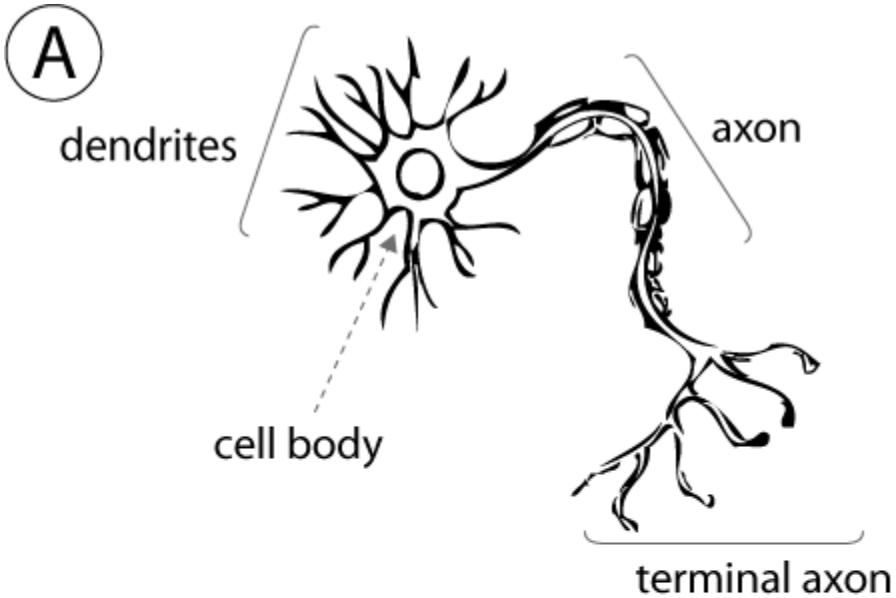
Ένα τεχνητό νευρωνικό δίκτυο είναι ένας συμπαγής παράλληλος κατανεμημένος επεξεργαστής που έχει τη φυσική κλίση να αποθηκεύει εμπεριστατωμένη γνώση και να την κάνει διαθέσιμη για χρήση.

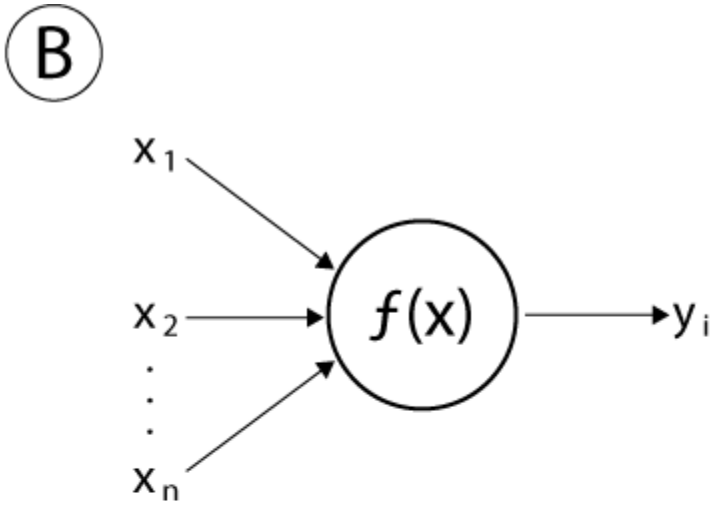
- Κύριο Πλεονέκτημα: παρουσιάζει ανοχή σε δεδομένα εκπαίδευσης με θόρυβο, δηλαδή δεδομένα που περιστασιακά έχουν λανθασμένες τιμές (π.χ. λάθη καταχώρησης).

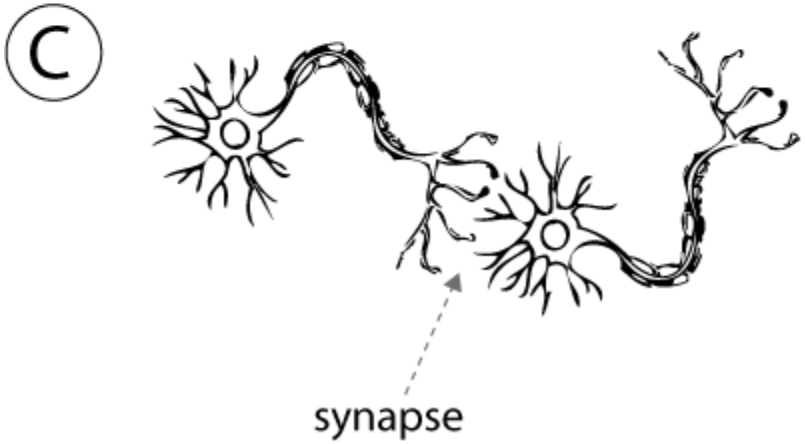
- Κύριο Μειονέκτημα: αδυνατεί να εξηγήσει ποιοτικά τη γνώση που μοντελοποιεί.

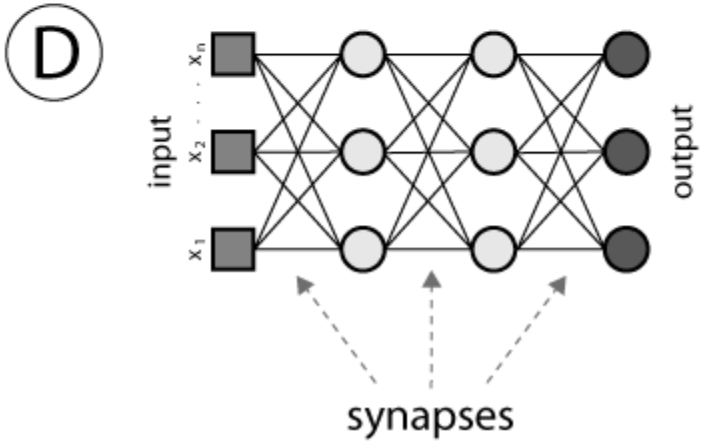


Νευρωνικά Δίκτυα









Νευρωνικά Δίκτυα και κλασικό λογισμικό

Ο τρόπος με τον οποίο πραγματοποιούνται οι διάφοροι υπολογισμοί μέσω Νευρωνικών Δικτύων διαφέρει ιδιαίτερα από αυτόν που χρησιμοποιείται στον κλασικό προγραμματισμό.

Η κύρια διαφορά τους είναι ότι το Νευρωνικό Δίκτυο μπορεί να εκπαιδευτεί χρησιμοποιώντας δείγματα (δεδομένα και αποτελέσματα) μιας διαδικασίας και να δημιουργήσει ένα τεχνητό μοντέλο αυτής αντί να προγραμματίζει το φυσικό μοντέλο της διαδικασίας το οποίο συνήθως δεν είναι γνωστό.

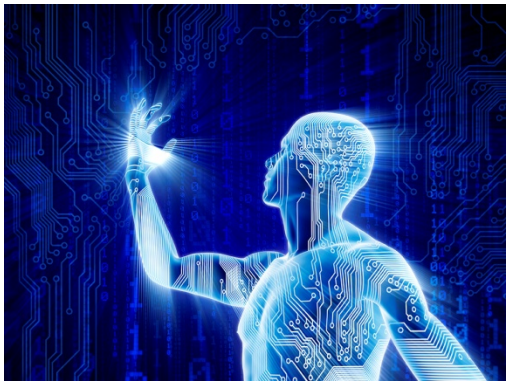
Νευρωνικά Δίκτυα	**Κλασικός προγραμματισμός**
1. Σύγχρονος τρόπος λειτουργίας	Ασύγχρονος τρόπος λειτουργίας
2. Παράλληλη επεξεργασία	Σειριακή επεξεργασία
3. Εκπαιδεύονται με παραδείγματα	Προγραμματίζεται
4. Συνύπαρξη μνήμης και μονάδων επεξεργασίας	Διαχωρισμός μνήμης και μονάδων επεξεργασίας
5. Ανοχή στα σφάλματα	Καμία ανοχή στα σφάλματα
6. Ο χρόνος ενός κύκλου είναι της τάξης του msec	Ο χρόνος ενός κύκλου είναι της τάξης του nsec

Εφαρμογές Τεχνητών Νευρωνικών Δικτύων

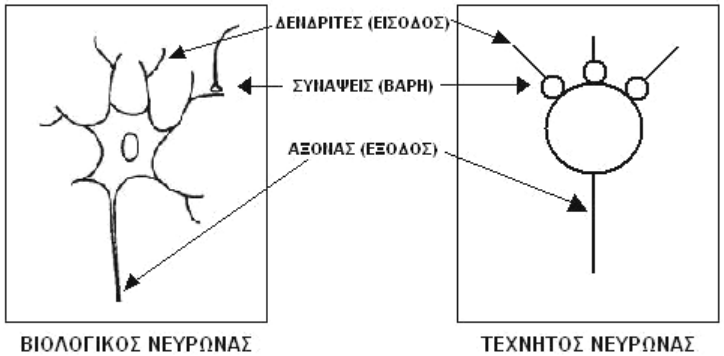
Τα ΤΝΔ προσφέρουν λύσεις κυρίως σε προβλήματα που σχετίζονται με τον ανθρώπινο παράγοντα, όπως αναγνώριση ομιλίας, εικόνας, κειμένου κλπ.

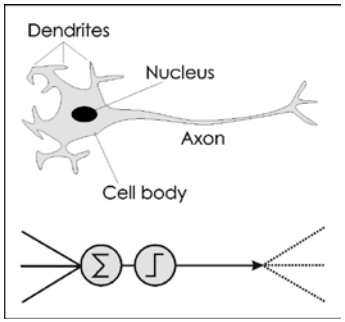
Σχεδιασμός ενεργειών, Χρονοπρογραμματισμός
Αναγνώριση προτύπων (π.χ. υπογραφών, ομιλίας, γραφής, εικόνων, κλπ)
Πρόβλεψη (π.χ. καιρού, τιμών μετοχών)
Συστήματα ελέγχου που βασίζονται στην ασαφή λογική

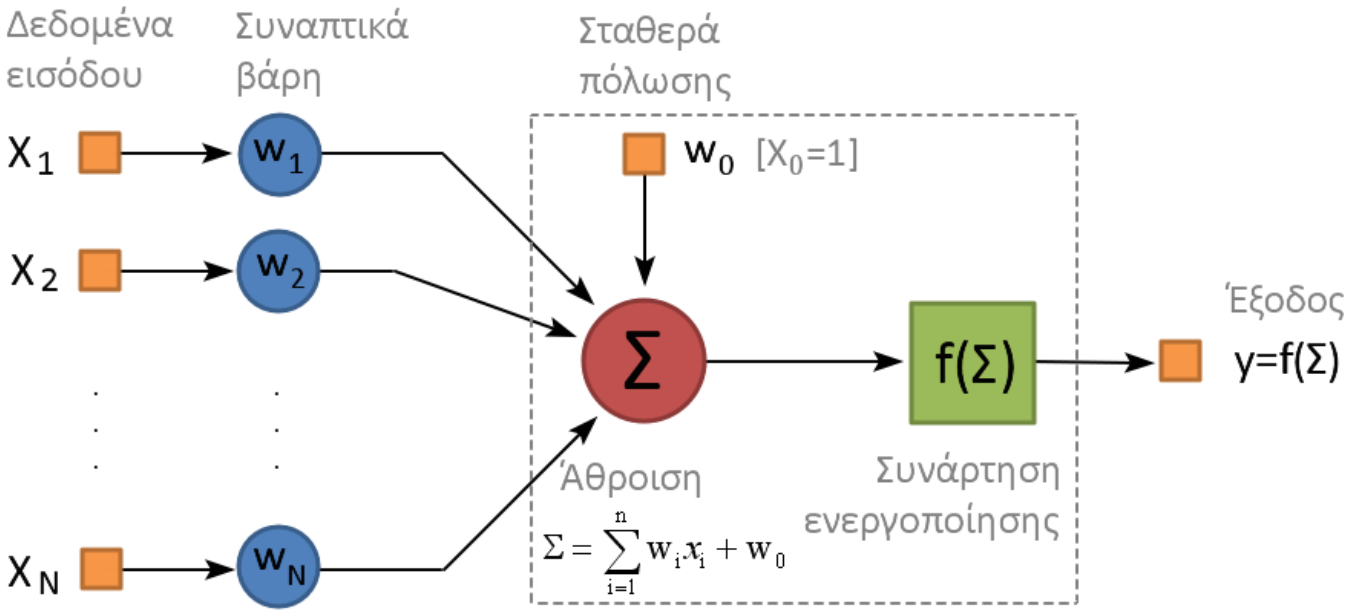
Βιολογία | Μοντέλα για την όραση
Ιατρική | Ανάγνωση και ανάλυση των ακτίνων Χ
Στρατιωτική τέχνη | Αναγνώριση και παρακολούθηση στόχων
Χρηματοοικονομικά | Ανάλυση επικινδυνότητας δανείων
Βιομηχανία | Έλεγχος στην γραμμή παραγωγής
Περιβάλλον | Πρόβλεψη καιρού



Στοιχειώδης Αισθητήρας – Elementary Perceptron



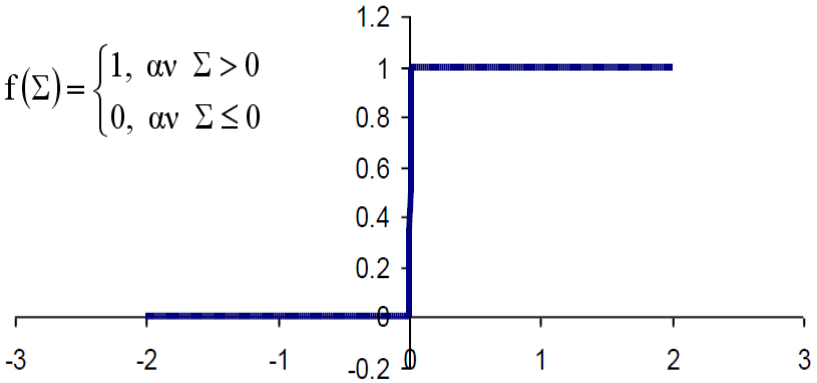




Συνάρτηση Ενεργοποίησης - Activation Function

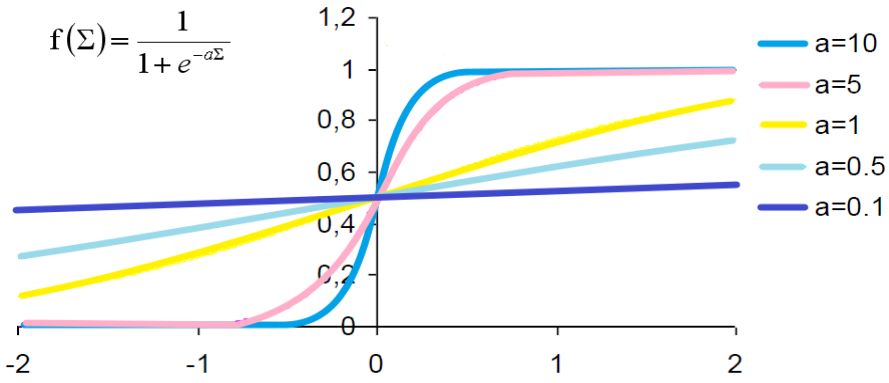
Βηματική συνάρτηση

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



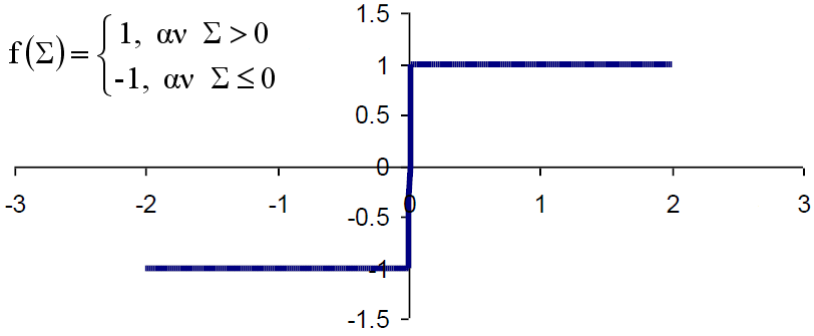
Σιγμοειδής συνάρτηση

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



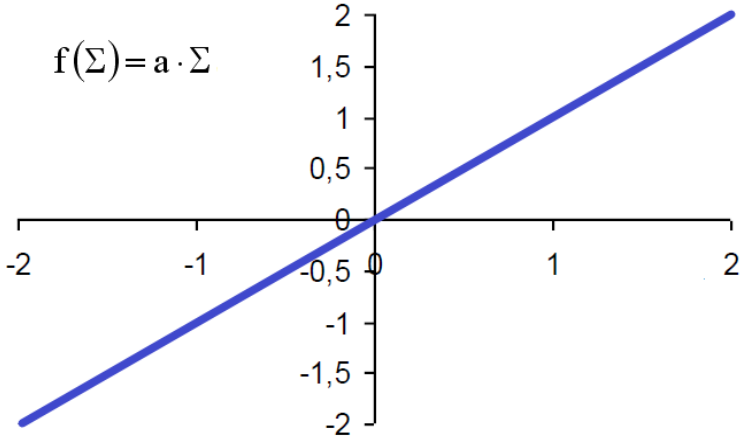
Συνάρτηση προσήμου

$$f(\Sigma)=\begin{cases}1, & \text{αν } \Sigma>0\\ -1, & \text{αν } \Sigma\le 0\end{cases}$$

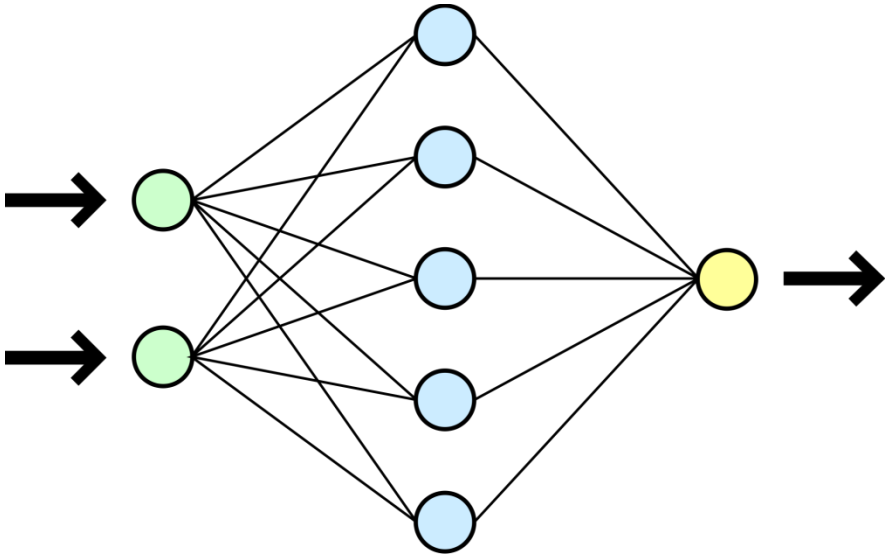


Γραμμική συνάρτηση ενεργοποίησης

$$f(\Sigma)=a\cdot\Sigma$$



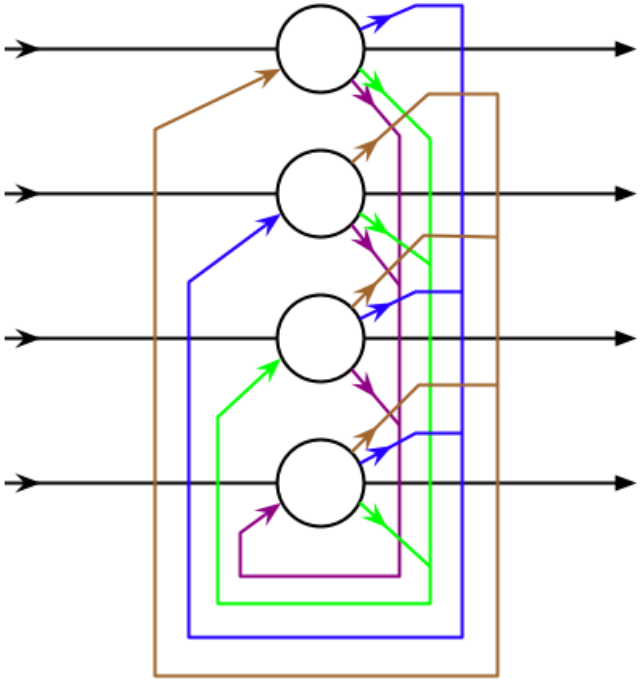
Τοπολογία Νευρωνικών Δικτύων



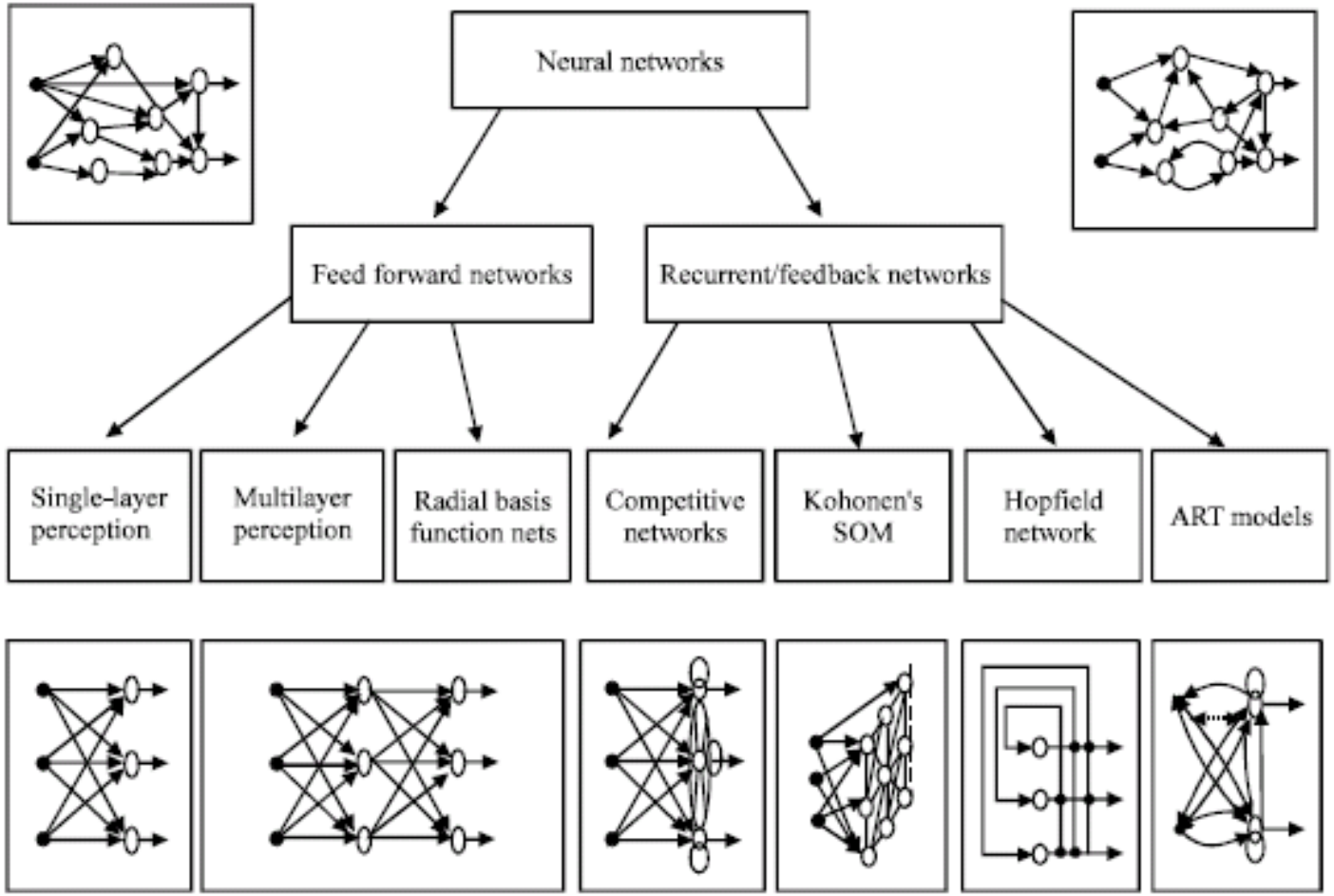
Δίκτυο με μετάδραση
(feedforward)

Δίκτυο με ανατροφοδότηση

(feedback)

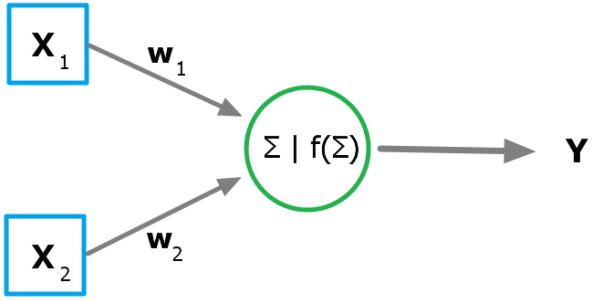


Τοπολογία Νευρωνικών Δικτύων



Διαδικασία Μάθησης

Τα ΤΝΔ πραγματοποιούν δύο βασικές λειτουργίες: τη μάθηση-εκπαίδευση (learning) και την ανάκληση (recall).



Μάθηση (εκπαίδευση) είναι η διαδικασία της τροποποίησης της τιμής των βαρών του δικτύου w_i ώστε, δοθέντος συγκεκριμένου (γνωστού) διανύσματος εισόδου, να παραχθεί συγκεκριμένο (γνωστό) διάνυσμα εξόδου (δημιουργία μοντέλου), η μάθηση πραγματοποιείται σε ένα μέρος των διαθεσίμων δεδομένων εισόδου- εξόδου (π.χ.70%).

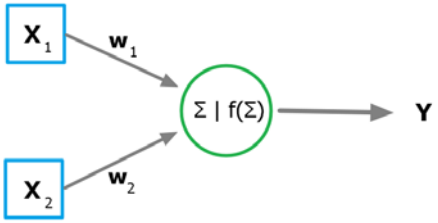
Ανάκληση είναι η διαδικασία του υπολογισμού ενός διανύσματος εξόδου για συγκεκριμένο διάνυσμα εισόδου με βάση τα βάρη που έχουν καθοριστεί από τη διαδικασία μάθησης (εφαρμογή μοντέλου, π.χ. για το υπόλοιπο 30%).

Διαδικασία μάθησης

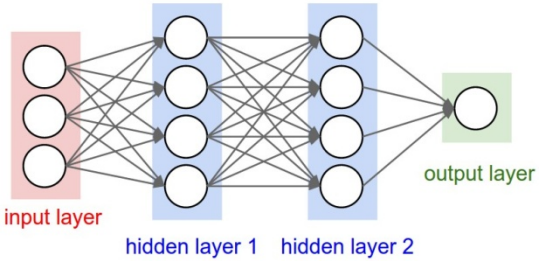
Μάθηση υπό επίβλεψη (supervised learning)
Στο δίκτυο δίνονται ζευγάρια διανυσμάτων εισόδου - επιθυμητής εξόδου.

Μάθηση χωρίς επίβλεψη (unsupervised learning)
Δίνονται μόνο δεδομένα εισόδου
Στόχος δεν είναι η εκτίμηση των αποτελεσμάτων από τα δεδομένο εισόδου αλλά η επεξεργασία των δεδομένων εισόδου.

Εκπαίδευση Νευρωνικών Δικτύων
α) Κανόνας Δέλτα (Delta rule)
ΤΝΔ τα οποία δεν έχουν κρυφά (ενδιάμεσα) επίπεδα



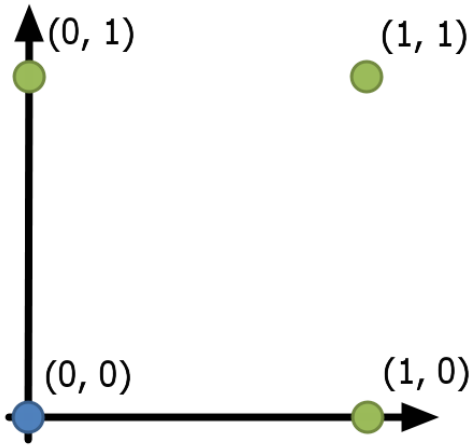
β) Ανάστροφη μετάδοση λάθους (Back propagation)
Αποτελεί την πιο γνωστή μέθοδο εκπαίδευσης νευρωνικών δικτύων πολλών επιπέδων.

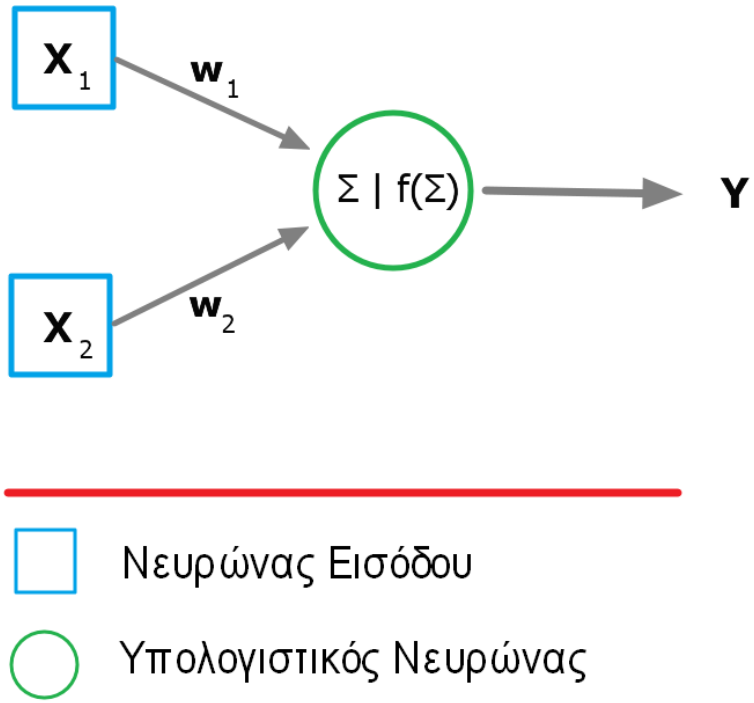


Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

Συνάρτηση OR

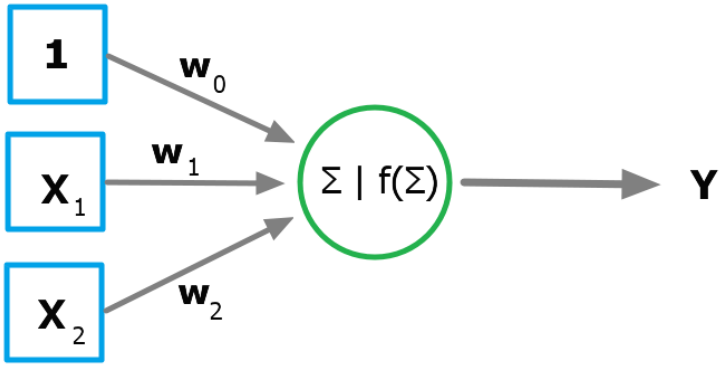
X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1





Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1



$$\Sigma = w_1 x_1 + w_2 x_2 + w_0$$

$$\mathrm{f}(\Sigma) = \Sigma$$

$$\mathrm{y} = \mathrm{f}(\Sigma) \rightarrow \mathrm{y} = w_1 x_1 + w_2 x_2 + w_0$$

$$\mathrm{w}_{1,\mathrm{new}} = w_1 + a \cdot \delta \cdot \mathrm{x}_1$$

$$\mathrm{w}_{2,\mathrm{new}} = w_2 + a \cdot \delta \cdot \mathrm{x}_2$$

$$\mathrm{w}_{0,\mathrm{new}} = w_0 + a \cdot \delta$$

$$\delta = \mathrm{t} - \mathrm{y}$$

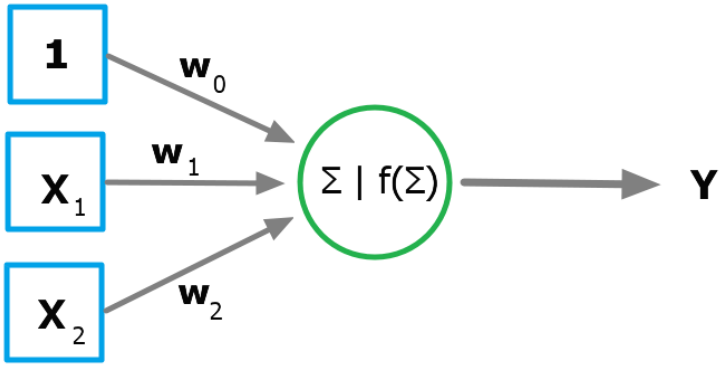
t: σωστή τιμή εξόδου

δ: σφάλμα (πρόβλεψης από τη σωστή τιμή εξόδου)

α: σταθερά εκμάθησης

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1



Βήμα 1

α =0,02

Έστω w_1=1, w_2=-1, w_0=0

για X_1=0 και X_2=0 → t =0

$$y(1) = w_1(1)\cdot x_1(1) + w_2(1)\cdot x_2 + w_0(1) = 1\cdot 0 + (-1)\cdot 0 + 0 = 0$$

$$t(1) = 0$$

$$\delta(1) = \mathrm{t}(1) - \mathrm{y}(1) = 0 - 0 = 0$$

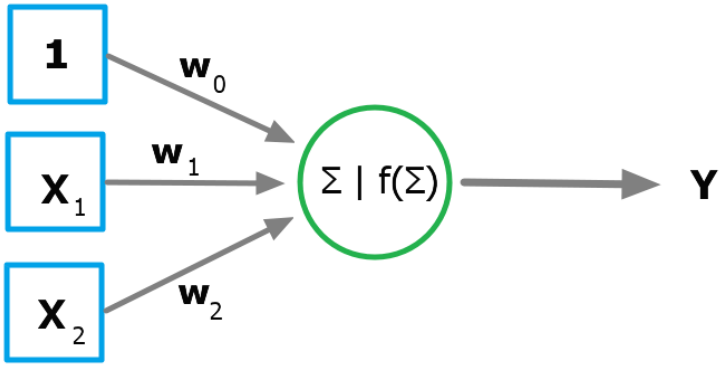
$$\mathrm{w}_1(2) = w_1(1) + a\cdot\delta(1)\cdot \mathrm{x}_1(1) = 1 + 0.02\cdot 0\cdot 0 = 1$$

$$\mathrm{w}_2(2) = w_2(1) + a\cdot\delta(1)\cdot \mathrm{x}_2(1) = -1 + 0.02\cdot 0\cdot 0 = -1$$

$$\mathrm{w}_0(2) = w_0(1) + a\cdot\delta(1) = 0 + 0.02\cdot 0 = 0$$

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1



Βήμα 2

για X_1=1 και X_2=0 → t =1

$$y(2) = w_1(2)\cdot x_1(2) + w_2(2)\cdot x_2 + w_0(2) = 1\cdot 1 + (-1)\cdot 0 + 0 = 1$$

$$t(2) = 1$$

$$\delta(2) = \mathrm{t}(2) - \mathrm{y}(2) = 1 - 1 = 0$$

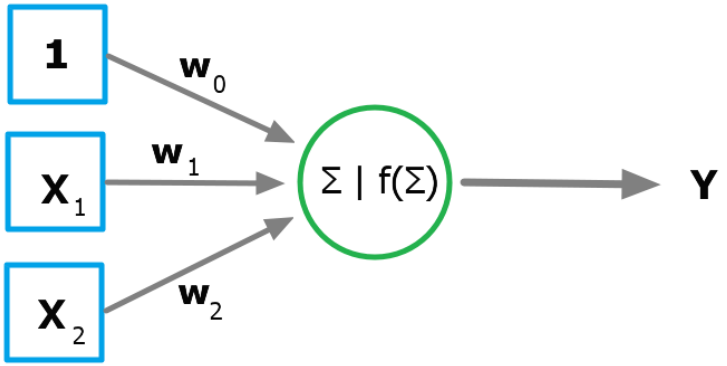
$$\mathrm{w}_1(3) = w_1(2) + a\cdot\delta(2)\cdot \mathrm{x}_1(2) = 1 + 0.02\cdot 0\cdot 1 = 1$$

$$\mathrm{w}_2(3) = w_2(2) + a\cdot\delta(2)\cdot \mathrm{x}_2(2) = -1 + 0.02\cdot 0\cdot 0 = -1$$

$$\mathrm{w}_0(3) = w_0(2) + a\cdot\delta(2) = 0 + 0.02\cdot 0 = 0$$

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1



Βήμα 3

για X_1=0 και X_2=1

→ t =1

$$y(3) = w_1(3)\cdot x_1(3) + w_2(3)\cdot x_2 + w_0(3) =$$

$$\rightarrow y(3) = 1\cdot 0 + (-1)\cdot 1 + 0 = -1$$

$$t(3) = 1$$

$$\delta(3) = \mathrm{t}(3) - \mathrm{y}(3) = 1 - (-1) = 2$$

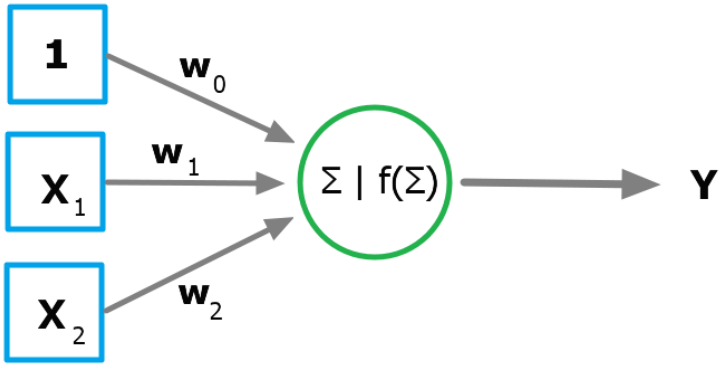
$$\mathrm{w}_1(4) = w_1(3) + a\cdot\delta(3)\cdot \mathrm{x}_1(3) = 1 + 0.02\cdot 2\cdot 0 = 1$$

$$\mathrm{w}_2(4) = w_2(3) + a\cdot\delta(3)\cdot \mathrm{x}_2(3) = -1 + 0.02\cdot 2\cdot 1 = -0.96$$

$$\mathrm{w}_0(4) = w_0(3) + a\cdot\delta(3) = 0 + 0.02\cdot 2 = 0.04$$

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1



Βήμα 4

για X_1=1 και X_2=1

→ t =1

$$y(4) = w_1(4)\cdot x_1(4) + w_2(4)\cdot x_2 + w_0(4) =$$

$$\rightarrow y(4) = 1\cdot 1 + (-0.96)\cdot 1 + 0.04 = 0.08$$

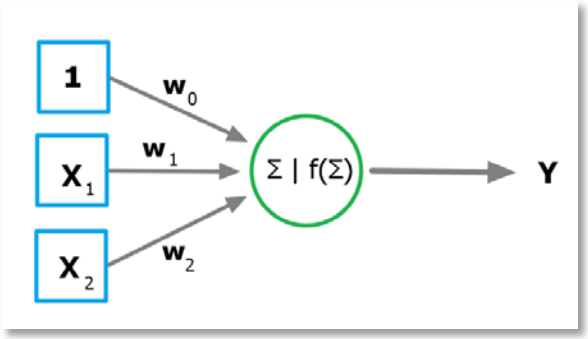
$$t(4) = 1$$

$$\delta(4) = \mathrm{t}(4) - \mathrm{y}(4) = 1 - 0.08 = 0.92$$

$$\mathrm{w}_1(5) = w_1(4) + a\cdot\delta(4)\cdot \mathrm{x}_1(4) = 1 + 0.02\cdot 0.92\cdot 1 = 1.02$$

$$\mathrm{w}_2(5) = w_2(4) + a\cdot\delta(4)\cdot \mathrm{x}_2(4) = -0.96 + 0.02\cdot 0.92\cdot 1 = -0.94$$

$$\mathrm{w}_0(5) = w_0(4) + a\cdot\delta(4) = 0.004 + 0.02\cdot 0.92 = 0.06$$



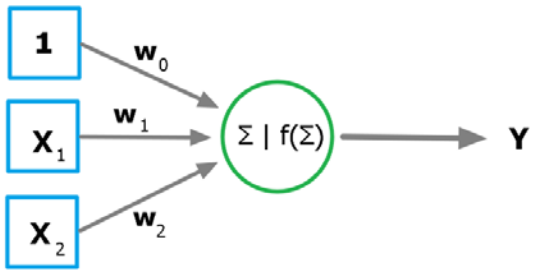
	Δεδομένα									new		
	x1	x2	t	w1	w2	w0	Σ	F(Σ)	δ	w1'	w2'	w0'
1	0	0	0	1.00	-1.00	0.00	0.00	0.00	0.00	1.00	-1.00	0.00
2	1	0	1	1.00	-1.00	0.00	1.00	1.00	0.00	1.00	-1.00	0.00
3	0	1	1	1.00	-1.00	0.00	-1.00	-1.00	2.00	1.00	-0.96	0.04
4	1	1	1	1.00	-0.96	0.04	0.08	0.08	0.92	1.02	-0.94	0.06

	Δεδομένα									new		
	x1	x2	t	w1	w2	w0	Σ	F(Σ)	δ	w1'	w2'	w0'
5	0	0	0	1.02	-0.94	0.06	0.06	0.06	-0.06	1.02	-0.94	0.06
6	1	0	1	1.02	-0.94	0.06	1.08	1.08	-0.08	1.02	-0.94	0.06
7	0	1	1	1.02	-0.94	0.06	-0.89	-0.89	1.89	1.02	-0.90	0.09
8	1	1	1	1.02	-0.90	0.09	0.21	0.21	0.79	1.03	-0.89	0.11

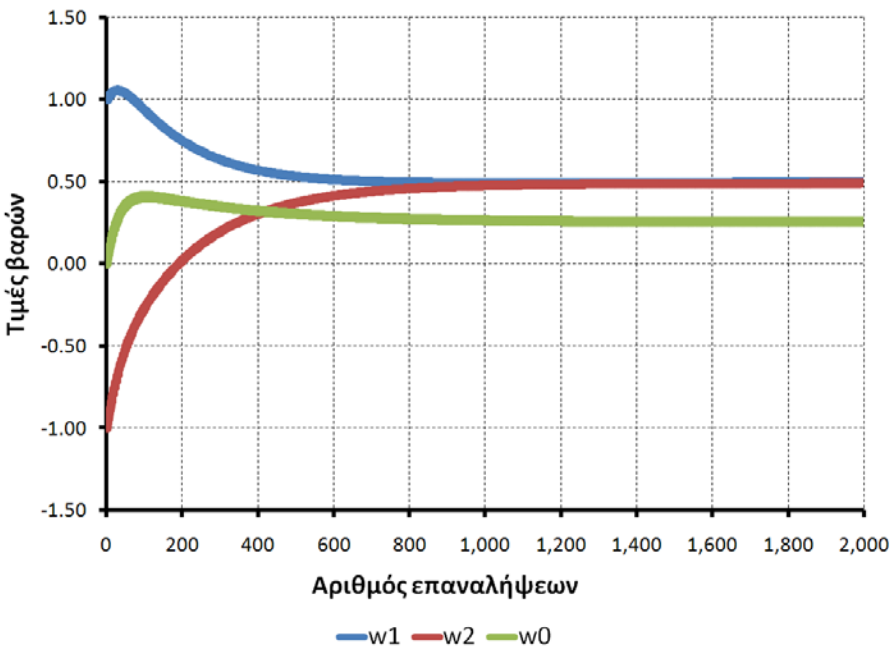
	Δεδομένα									new		
	x1	x2	t	w1	w2	w0	Σ	F(Σ)	δ	w1'	w2'	w0'
1997	0	0	0	0.49	0.49	0.26	0.26	0.26	-0.26	0.49	0.49	0.25
1998	1	0	1	0.49	0.49	0.25	0.75	0.75	0.25	0.50	0.49	0.26
1999	0	1	1	0.50	0.49	0.26	0.74	0.74	0.26	0.50	0.49	0.26
2000	1	1	1	0.50	0.49	0.26	1.25	1.25	-0.25	0.49	0.49	0.26
	Μέσος όρος (w)			0.50	0.49	0.26						

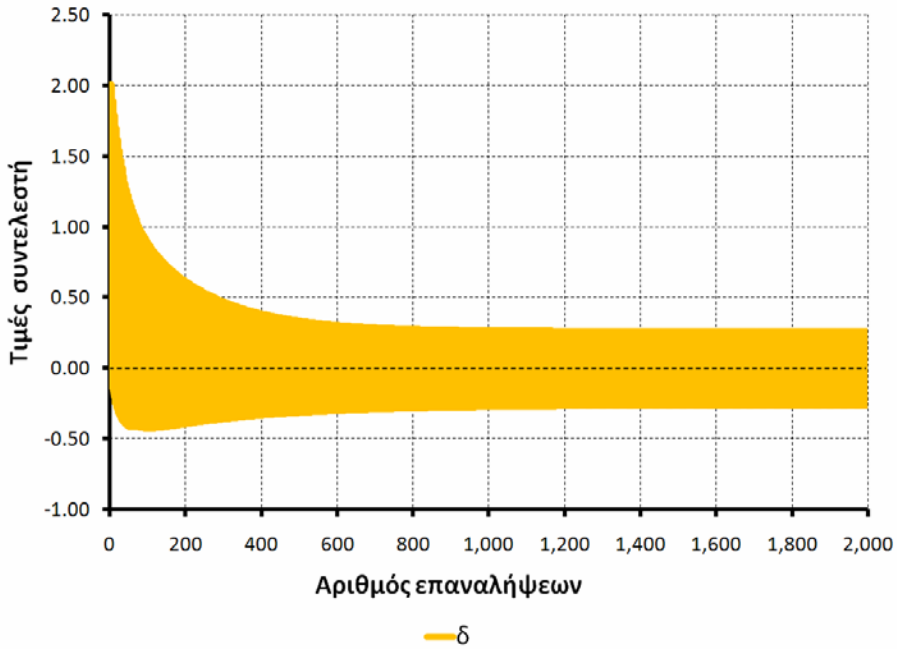
$$y = 0.50x_1 + 0.49x_2 + 0.26$$

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)



$$y = 0.50x_1 + 0.49x_2 + 0.26$$

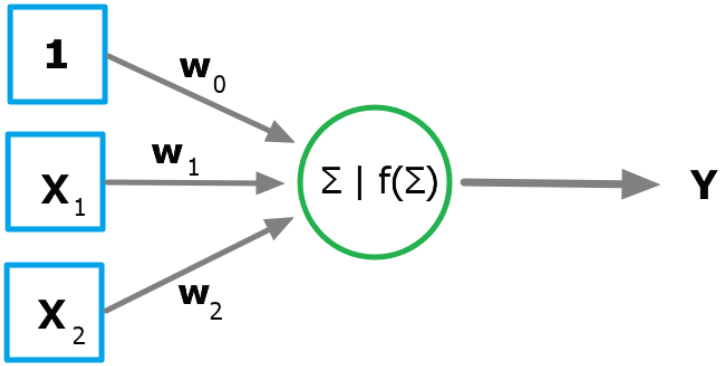




Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

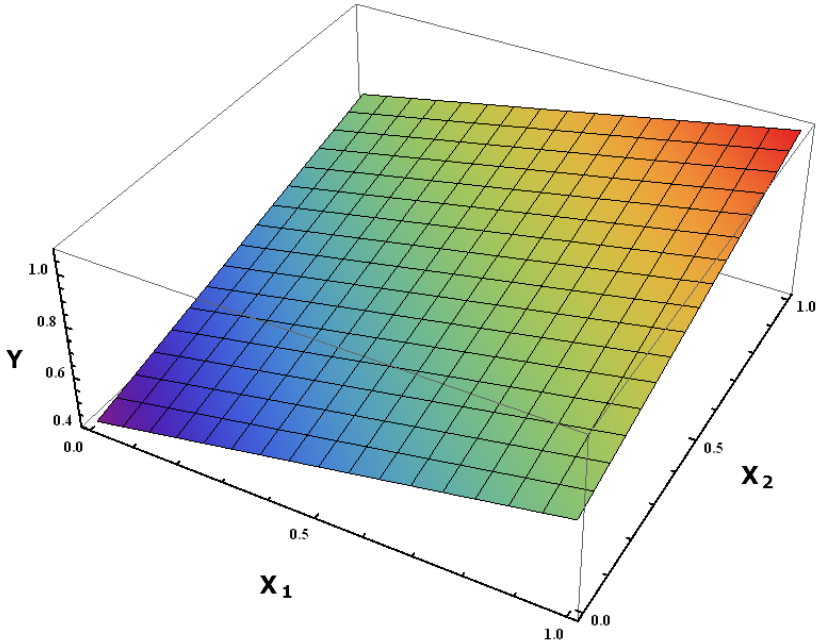
Αποτελέσματα από νευρωνικό δίκτυο

$$y = 0.50x_1 + 0.49x_2 + 0.26$$

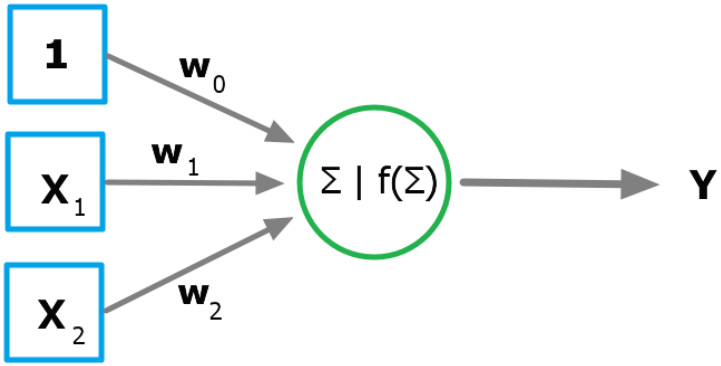


Αποτελέσματα

X1	X2	Y
0	0	0.26
1	0	0.75
0	1	0.75
1	1	1.24

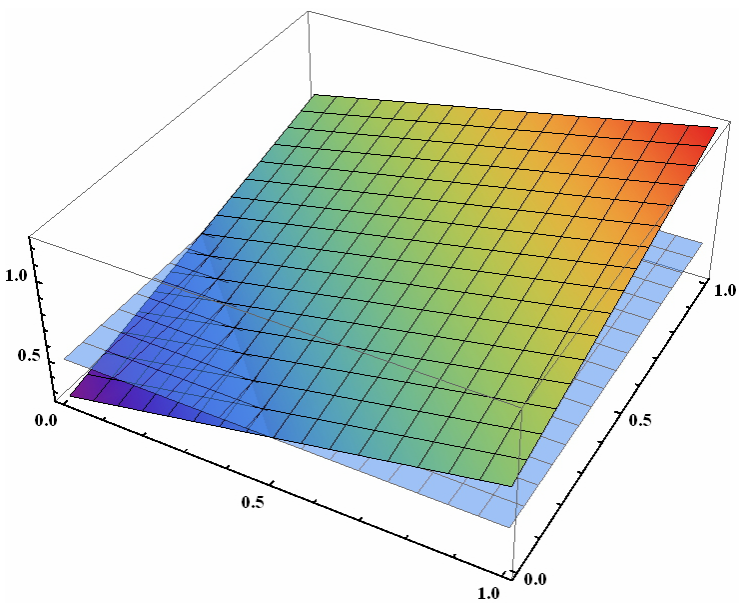


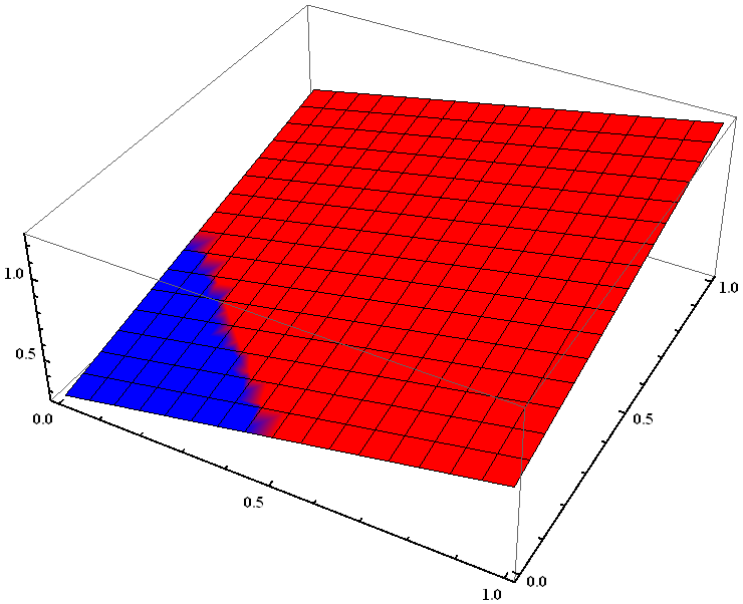
Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)



Αποτελέσματα

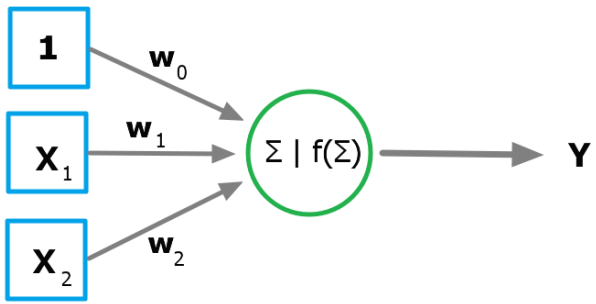
X1	X2	Y
0	0	0.26
1	0	0.75
0	1	0.75
1	1	1.24





Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1

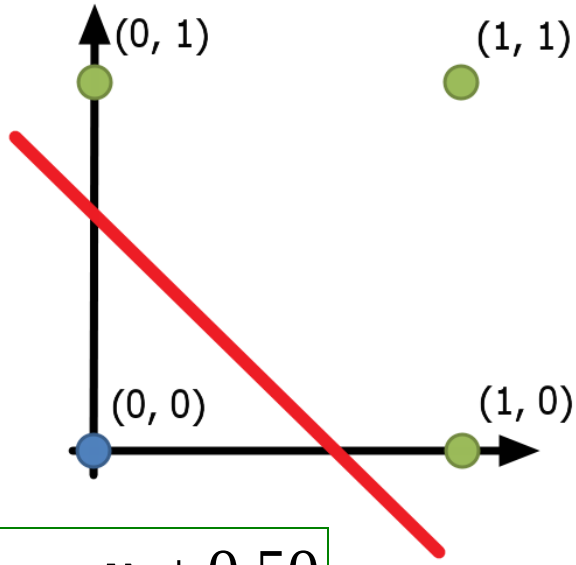


$$\Sigma = w_1 x_1 + w_2 x_2 + w_0$$

$$f(\Sigma) = \Sigma$$

$$y = f(\Sigma) \rightarrow$$

$$y = w_1 x_1 + w_2 x_2 + w_0$$



$$x_2 = -x_1 + 0.50$$

$$y = w_1 x_1 + w_2 x_2 + w_0$$

$$\acute{\varepsilon}\sigma\tau\omega \quad y = 0.5$$

$$0.5 = w_1 x_1 + w_2 x_2 + w_0$$

$$x_2 = -\frac{w_1}{w_2} x_1 + \frac{0.5 - w_0}{w_2}$$

εύρεση ευθείας ορίου απόφασης

Παράδειγμα 1 (ΤΝΔ vs. Παλινδρόμηση)

Πολλαπλή γραμμική παλινδρόμηση

$$y = 0.5x_1 + 0.5x_2 + 0.25$$

$$R = 0.82$$

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1

Αποτελέσματα παλινδρόμησης

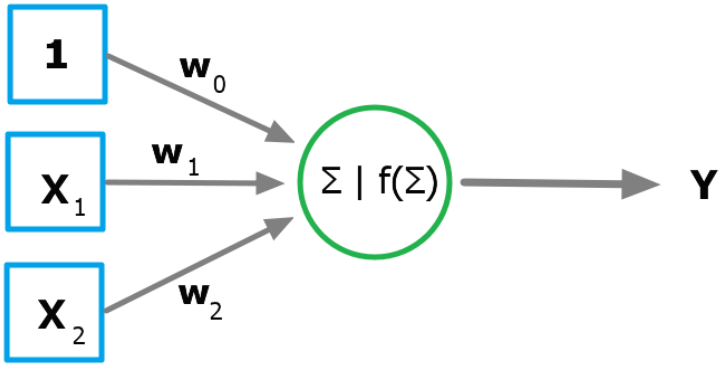
X1	X2	Y
0	0	0.25
1	0	0.50
0	1	0.50
1	1	1.00

Αποτελέσματα ΤΝΔ

X1	X2	Y
0	0	0.26
1	0	0.75
0	1	0.75
1	1	1.24

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	1



$$\Sigma = w_1 x_1 + w_2 x_2 + w_0$$

$$\mathrm{f}(\Sigma) = \frac{1}{1+e^{-10\cdot\Sigma}}$$

$$\mathrm{y} = \mathrm{f}(\Sigma) \rightarrow \mathrm{y} = \frac{1}{1+e^{-10\cdot(w_1 x_1 + w_2 x_2 + w_0)}}$$

$$\delta = \mathrm{t} - \mathrm{y}$$

t: σωστή τιμή εξόδου

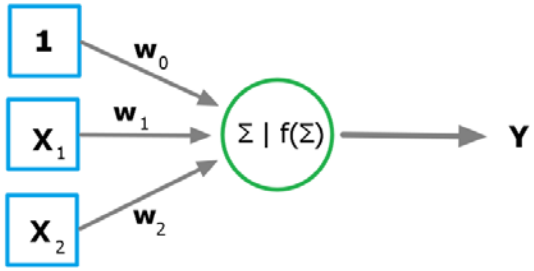
δ: σφάλμα (πρόβλεψης από τη σωστή τιμή εξόδου)

$$\mathrm{w}_{1,\mathrm{new}} = w_1 + a \cdot \delta \cdot \mathrm{x}_1$$

$$\mathrm{w}_{2,\mathrm{new}} = w_2 + a \cdot \delta \cdot \mathrm{x}_2$$

$$\mathrm{w}_{0,\mathrm{new}} = w_0 + a \cdot \delta$$

α: σταθερά εκμάθησης



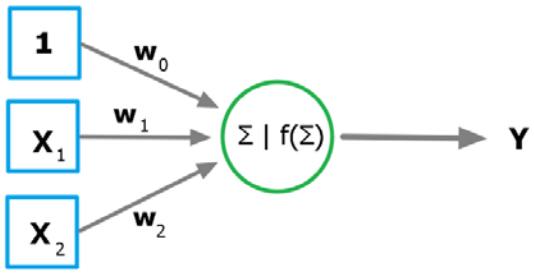
	Δεδομένα									new		
	x1	x2	t	w1	w2	w0	Σ	F(Σ)	δ	w1'	w2'	w0'
1	0	0	0	1.00	-1.00	0.00	0.00	0.50	-0.50	1.00	-1.00	-0.05
2	1	0	1	1.00	-1.00	-0.05	0.95	0.99	0.01	1.00	-1.00	-0.05
3	0	1	1	1.00	-1.00	-0.05	-1.05	0.01	0.99	1.00	-0.90	0.05
4	1	1	1	1.00	-0.90	0.05	0.15	0.68	0.32	1.03	-0.87	0.08

	Δεδομένα									new		
	x1	x2	t	w1	w2	w0	Σ	F(Σ)	δ	w1'	w2'	w0'
5	0	0	0	1.03	-0.87	0.08	0.08	0.60	-0.60	1.03	-0.87	0.02
6	1	0	1	1.03	-0.87	0.02	1.06	0.99	0.01	1.03	-0.87	0.02
7	0	1	1	1.03	-0.87	0.02	-0.85	0.01	0.99	1.03	-0.77	0.12
8	1	1	1	1.03	-0.77	0.12	0.38	0.87	0.13	1.05	-0.76	0.13

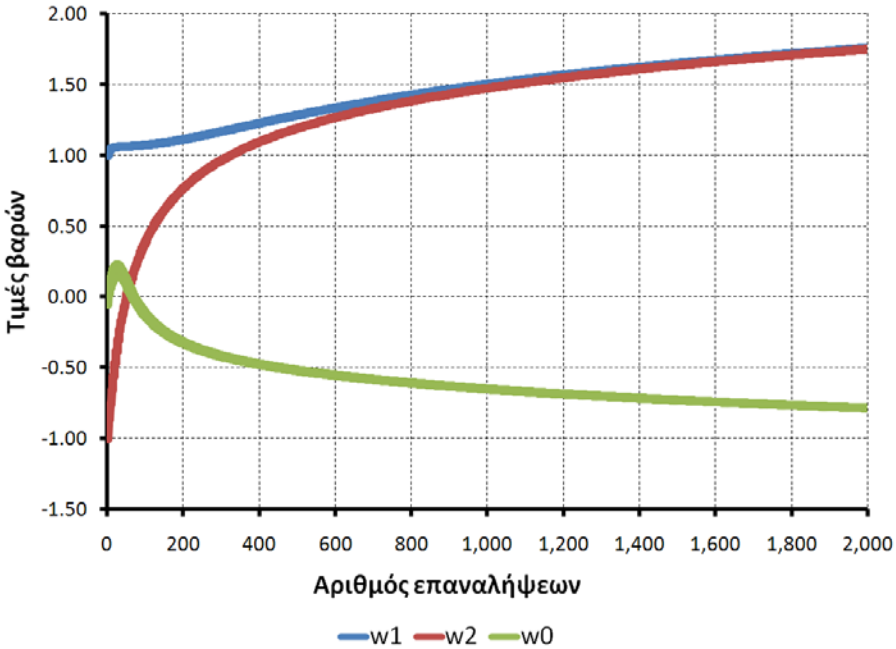
$$y = \frac{1}{1 + e^{-10\cdot(1.76\cdot x_1 + 1.75\cdot x_2 - 0.78)}}$$

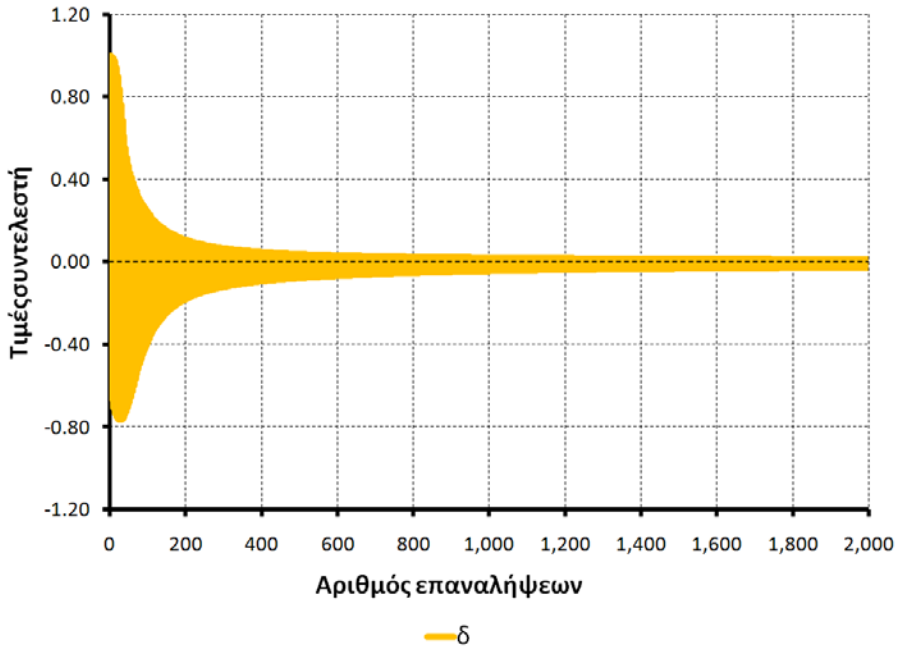
	Δεδομένα									new		
	x1	x2	t	w1	w2	w0	Σ	F(Σ)	δ	w1'	w2'	w0'
1997	0	0	0	1.76	1.75	-0.78	-0.78	0.02	-0.02	1.76	1.75	-0.79
1998	1	0	1	1.76	1.75	-0.79	0.97	0.99	0.01	1.76	1.75	-0.78
1999	0	1	1	1.76	1.75	-0.78	0.97	0.99	0.01	1.76	1.75	-0.78
2000	1	1	1	1.76	1.75	-0.78	2.73	1.00	0.00	1.76	1.75	-0.78
	Μέσος όρος (w)			1.76	1.75	-0.78						

Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)



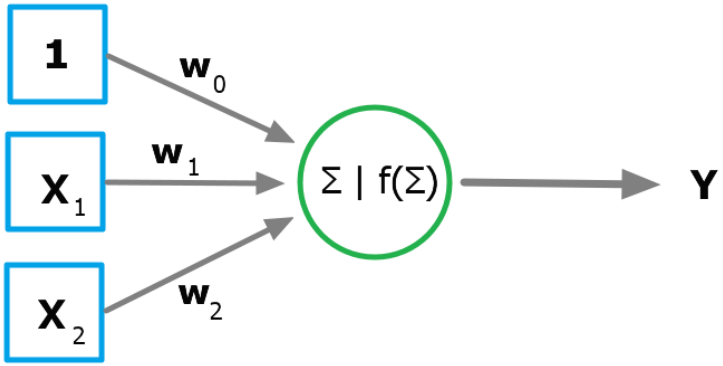
$$y = \frac{1}{1 + e^{-10\cdot(1.76\cdot x_1 + 1.75\cdot x_2 - 0.78)}}$$





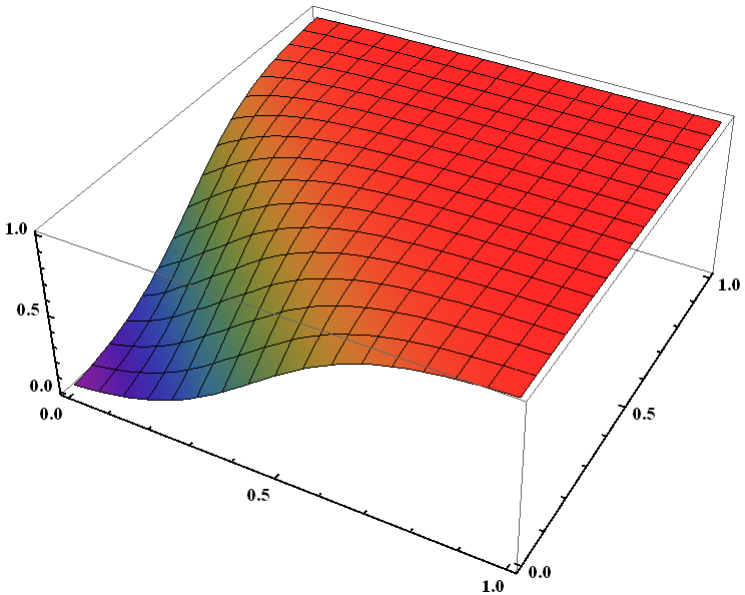
Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)

$$y = \frac{1}{1 + e^{-10\cdot(1.76\cdot x_1 + 1.75\cdot x_2 - 0.78)}}$$

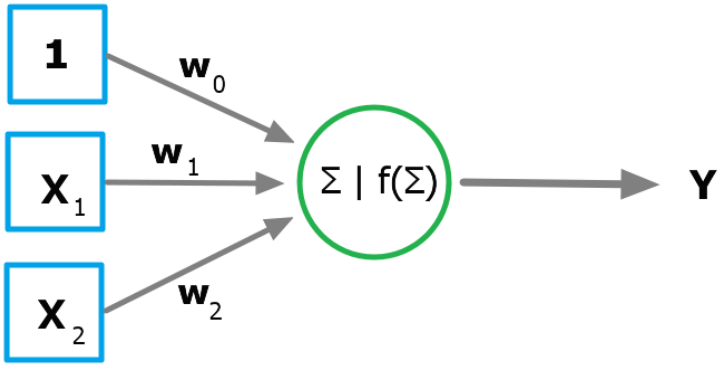


Αποτελέσματα

X1	X2	Y
0	0	0.02
1	0	0.99
0	1	0.99
1	1	1.00

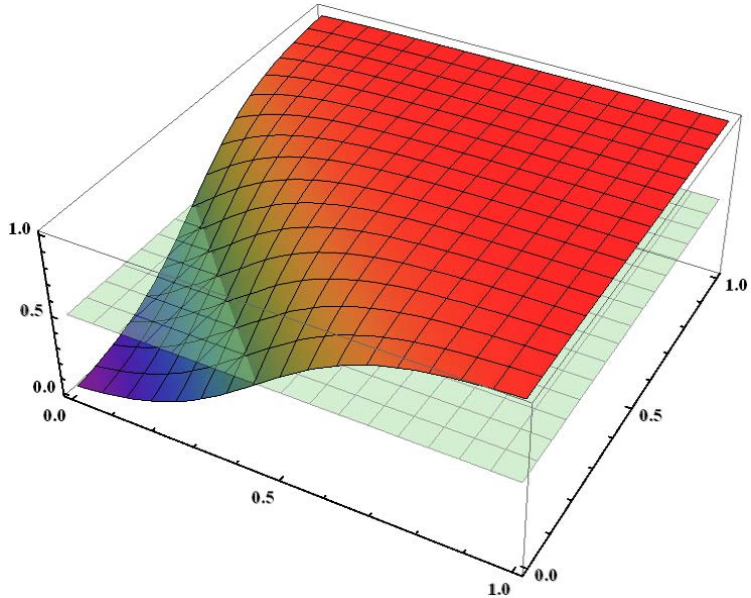


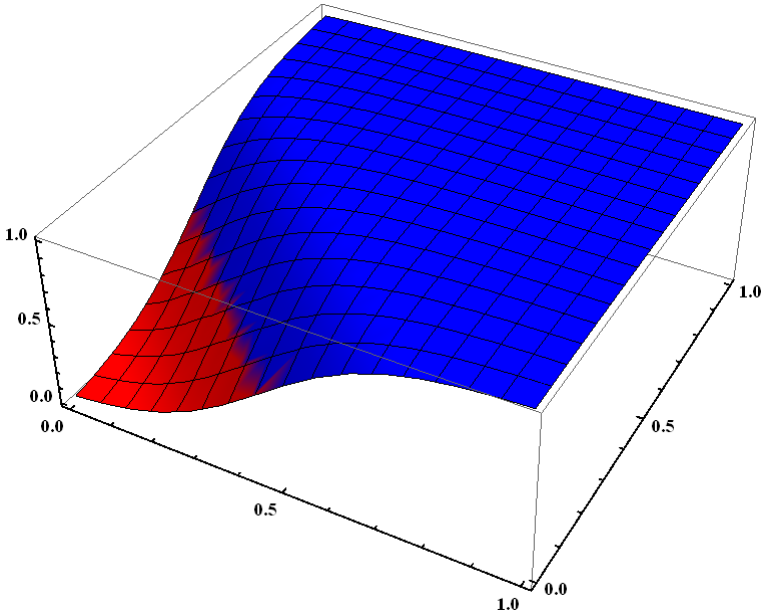
Παράδειγμα 1 (ΤΝΔ ενός επιπέδου)



Αποτελέσματα

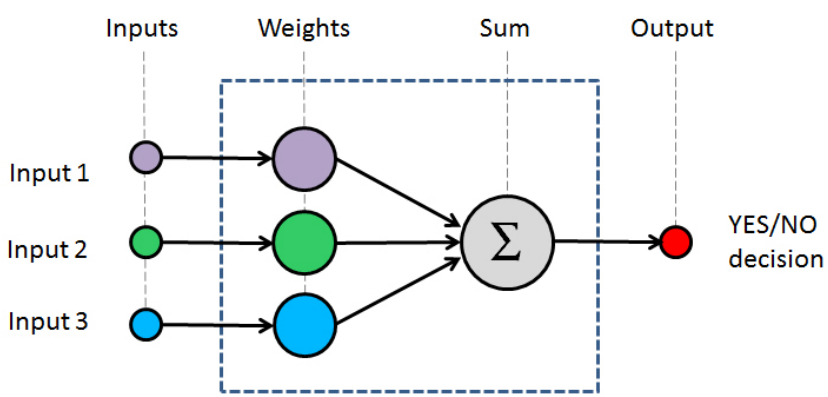
X1	X2	Y
0	0	0.02
1	0	0.99
0	1	0.99
1	1	1.00

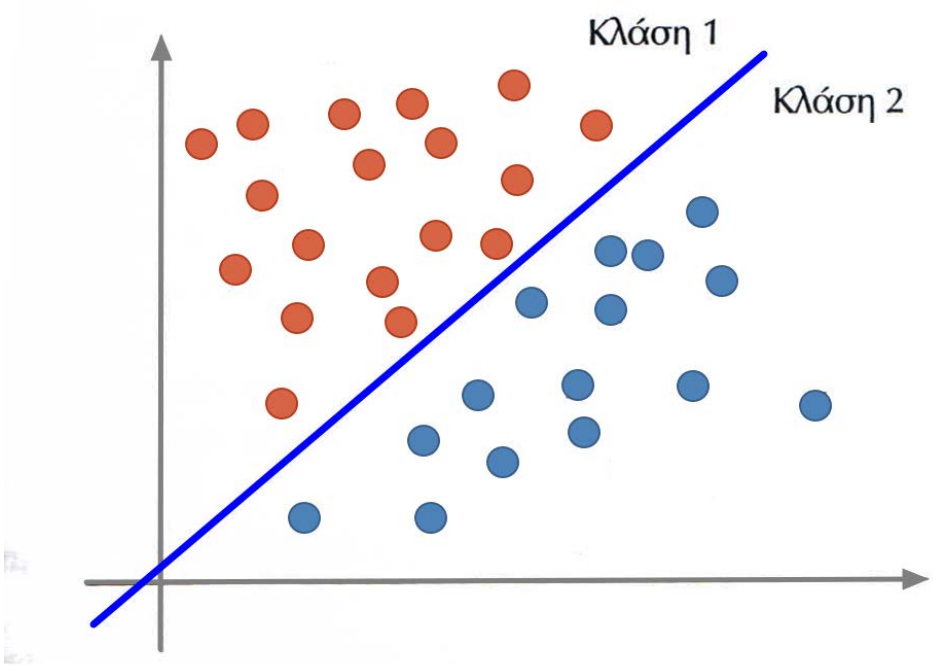




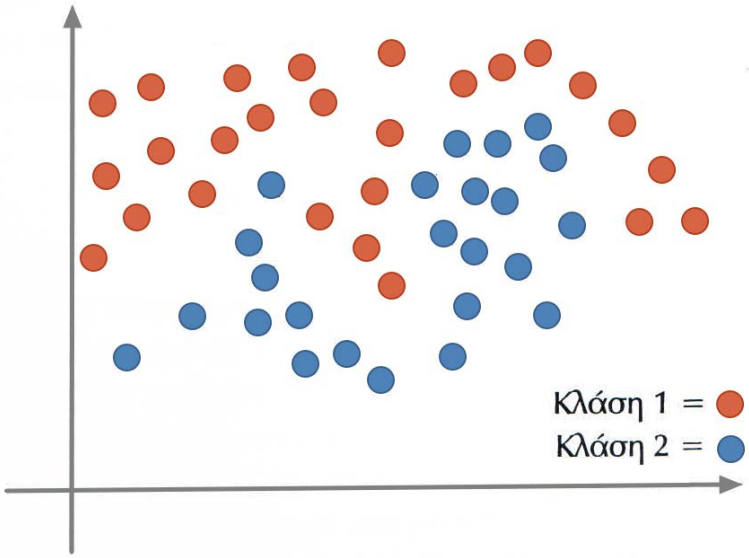
Νευρωνικά Δίκτυα - Γραμμικά ανεξάρτητες συναρτήσεις

Ένα απλό νευρωνικό δίκτυο ορίζει μια ευθεία που χωρίζει το επίπεδο σε δυο τμήματα δημιουργώντας ένα όριο απόφασης μεταξύ των δυο εναλλακτικών αποφάσεων ανάλογα με τη θέση των δεδομένων εισόδου (x1, x2), πάνω ή κάτω από την ευθεία γραμμή που χωρίζει τις δύο κλάσεις καλύτερα.

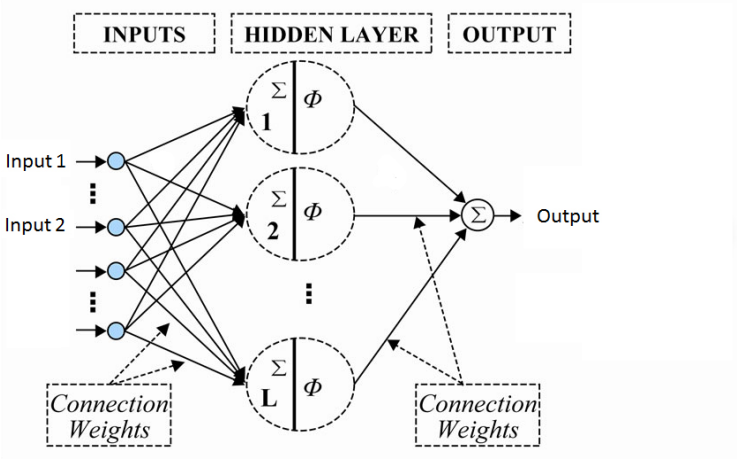


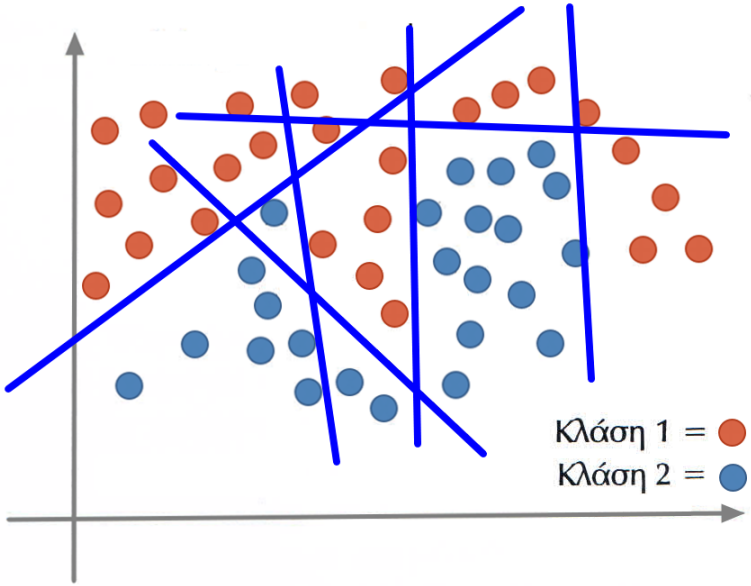


Νευρωνικά Δίκτυα – Μη γραμμικά ανεξάρτητες συναρτήσεις



Για μη γραμμικά ανεξάρτητα δεδομένα, χρησιμοποιούνται πολλαπλά όρια αποφάσεων και, κατ΄ επέκταση, πολυεπιπεδα Νευρωνικά Δίκτυα (Perceptron)

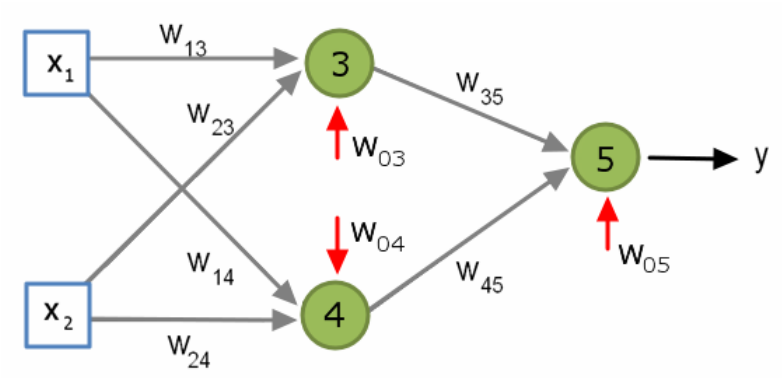


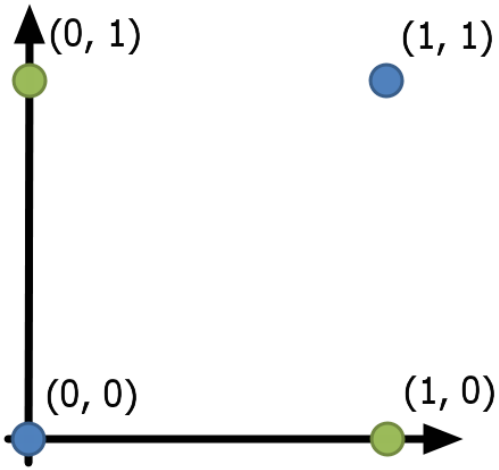


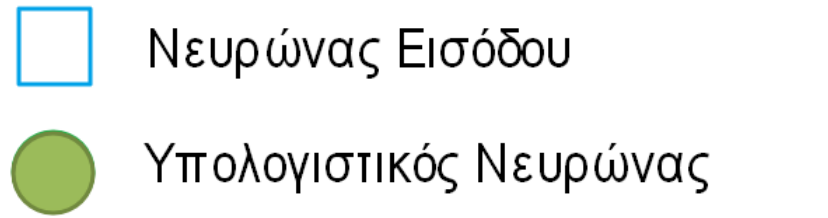
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

Συνάρτηση XOR

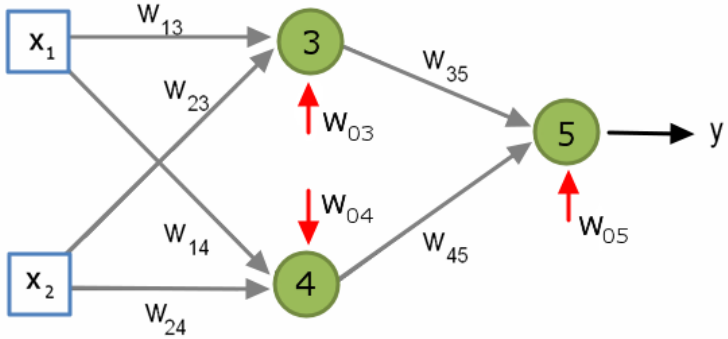
X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	0







Αλγόριθμος ανάστροφης μετάδοσης λάθους (Back propagation)



$$\Sigma_3 = w_{13}x_1 + w_{23}x_2 + w_{03}$$

$$\rightarrow y_3 = f(\Sigma_3) = \frac{1}{1+e^{-\Sigma_3}}$$

$$\Sigma_4 = w_{14}x_1 + w_{24}x_2 + w_{04}$$

$$\rightarrow y_4 = f(\Sigma_4) = \frac{1}{1+e^{-\Sigma_4}}$$

$$\Sigma_5 = w_{35}y_3 + w_{45}y_4 + w_{05}$$

$$\rightarrow y_5 = f(\Sigma_5) = \frac{1}{1+e^{-\Sigma_5}}$$

$$\delta_5 = (t - y_5) f'(\Sigma_5) = (t - y_5) \cdot f(\Sigma_5) \cdot (1 - f(\Sigma_5))$$

$$\delta_4 = w_{45}\delta_5 f'(\Sigma_4) = w_{45} \cdot \delta_5 \cdot f(\Sigma_4) \cdot (1 - f(\Sigma_4))$$

$$\delta_3 = w_{35}\delta_5 f'(\Sigma_3) = w_{35} \cdot \delta_5 \cdot f(\Sigma_3) \cdot (1 - f(\Sigma_3))$$

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

$$w_{13,new} = w_{13} + \delta_3 \cdot x_1$$

$$w_{23,new} = w_{23} + \delta_3 \cdot x_2$$

$$w_{14,new} = w_{14} + \delta_4 \cdot x_1$$

$$w_{24,new} = w_{24} + \delta_4 \cdot x_2$$

$$w_{35,new} = w_{35} + \delta_5 \cdot y_3$$

$$w_{45,new} = w_{45} + \delta_5 \cdot y_4$$

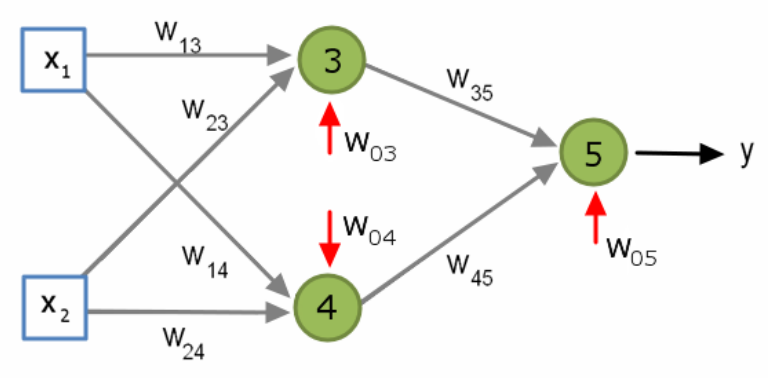
$$w_{03,new} = w_{03} + \delta_3$$

$$w_{04,new} = w_{04} + \delta_4$$

$$w_{05,new} = w_{05} + \delta_5$$

Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	0



Βήμα 1

Έστω w13=w14=w23=w24=w45=w35=1,
w03=1, w04=-1, w05=1

α=1

για X_1=0 και X_2=0 → t =0

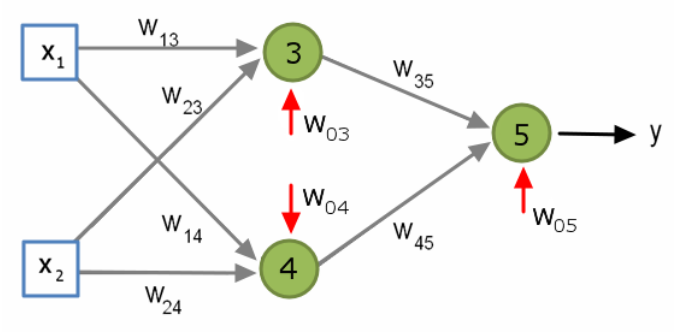
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

Βήμα 1

Έστω w13=w14=w23=w24=w45=w35=1, w03=1, w04=-1, w05=1

α=1

για X_1=0 και X_2=0 → t =0



$$\Sigma_3 = w_{13}x_1 + w_{23}x_2 + w_{03} = 1\cdot 0 + 1\cdot 0 + 1 = 1$$

$$\rightarrow y_3 = f(1) = \frac{1}{1+e^{-\Sigma_3}} = 0.73$$

$$\Sigma_4 = w_{14}x_1 + w_{24}x_2 + w_{04} = 1\cdot 0 + 1\cdot 0 - 1 = -1$$

$$\rightarrow y_4 = f(-1) = \frac{1}{1+e^{-\Sigma_4}} = 0.27$$

$$\Sigma_5 = w_{35}y_3 + w_{45}y_4 + w_{05} = -1\cdot 0.73 + 1\cdot 0.27 + 1.0 = 2$$

$$\rightarrow y_5 = f(2) = \frac{1}{1+e^{-\Sigma_5}} = 0.88$$

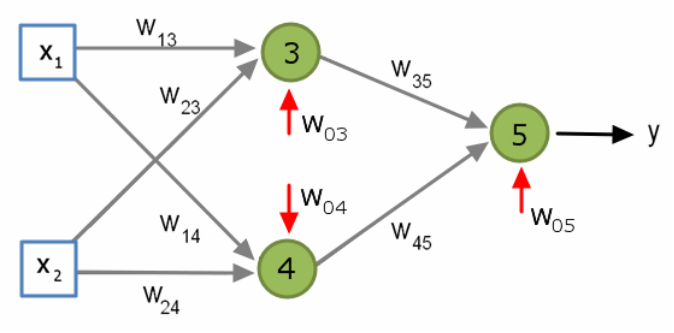
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

Βήμα 1

Έστω w13=w14=w23=w24=w45=w35=1, w03=1, w04=-1, w05=1

α=1

για X_1=0 και X_2=0 → t =0



$$\Sigma_3(1) = +1.00 \rightarrow y_3(1) = f(+1.00) = 0.73$$
$$\Sigma_4(1) = -1.00 \rightarrow y_4(1) = f(-1.00) = 0.27$$
$$\Sigma_5(1) = +2.00 \rightarrow y_5(1) = f(+2.00) = 0.88$$

$$\delta_5 = (t - y_5)\, f'(\Sigma_5) = (t - y_5) \cdot f(\Sigma_5) \cdot (1 - f(\Sigma_5)) = -0.09$$
$$\delta_4 = w_{45}\delta_5\, f'(\Sigma_4) = w_{45} \cdot \delta_5 \cdot f(\Sigma_4) \cdot (1 - f(\Sigma_4)) = -0.02$$
$$\delta_3 = w_{35}\delta_5\, f'(\Sigma_{34}) = w_{35} \cdot \delta_5 \cdot f(\Sigma_3) \cdot (1 - f(\Sigma_3)) = -0.02$$

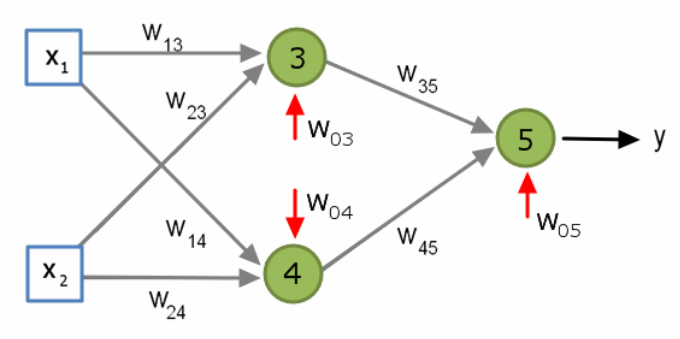
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

Βήμα 1

Έστω w13=w14=w23=w24=w45=w35=1, w03=1, w04=-1, w05=1

α=1

για X_1=0 και X_2=0 → t =0



$$\Sigma_3(1) = +1.00 \rightarrow y_3(1) = f(+1.00) = 0.73$$
$$\Sigma_4(1) = -1.00 \rightarrow y_4(1) = f(-1.00) = 0.27$$
$$\Sigma_5(1) = +2.00 \rightarrow y_5(1) = f(+2.00) = 0.88$$

$$\delta_5(1) = -0.09$$
$$\delta_4(1) = -0.02$$
$$\delta_3(1) = -0.02$$

$$w_{13,new} = w_{13} + \delta_3 \cdot x_1 = +1.00$$
$$w_{23,new} = w_{23} + \delta_3 \cdot x_2 = +1.00$$
$$w_{14,new} = w_{14} + \delta_4 \cdot x_1 = +1.00$$
$$w_{24,new} = w_{24} + \delta_4 \cdot x_2 = +1.00$$
$$w_{35,new} = w_{35} + \delta_5 \cdot y_3 = +0.93$$
$$w_{45,new} = w_{45} + \delta_5 \cdot y_4 = +0.98$$
$$w_{03,new} = w_{03} + \delta_3 = +0.98$$
$$w_{04,new} = w_{04} + \delta_4 = -1.02$$
$$w_{05,new} = w_{05} + \delta_5 = +0.91$$

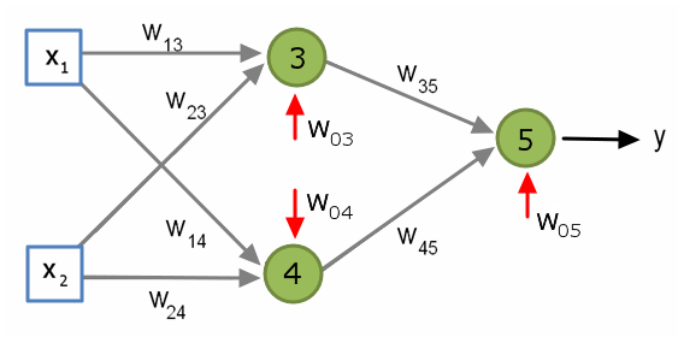
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

Βήμα 2

Έστω w13=w14=w23=w24=w45=w35=1, w03=1, w04=-1, w05=1

α=1

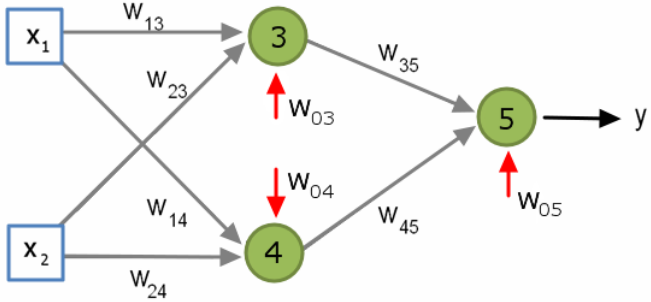
για X_1=1 και X_2=0 → t =1



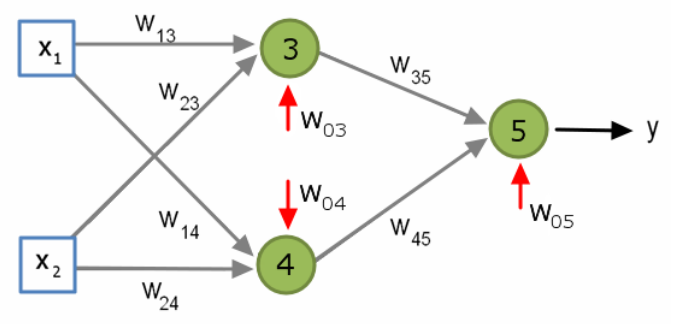
$$\Sigma_3(2) = +1.98 \rightarrow y_3(2) = f(+1.98) = 0.88$$
$$\Sigma_4(2) = -0.02 \rightarrow y_4(2) = f(-0.02) = 0.50$$
$$\Sigma_5(2) = +2.21 \rightarrow y_5(2) = f(+2.21) = 0.90$$

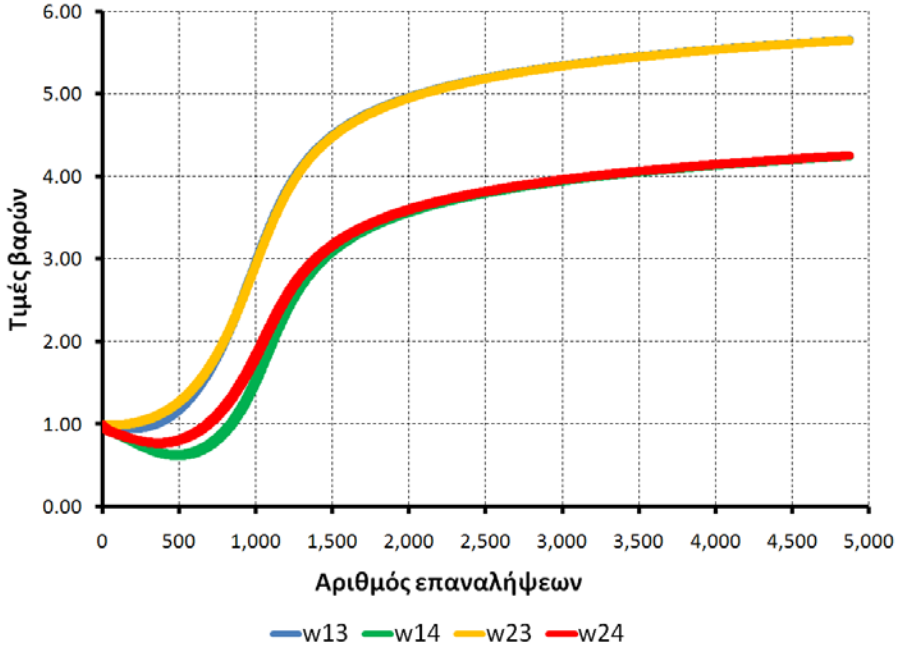
$$\delta_5(1) = +0.01$$
$$\delta_4(1) = +0.00$$
$$\delta_3(1) = +0.00$$

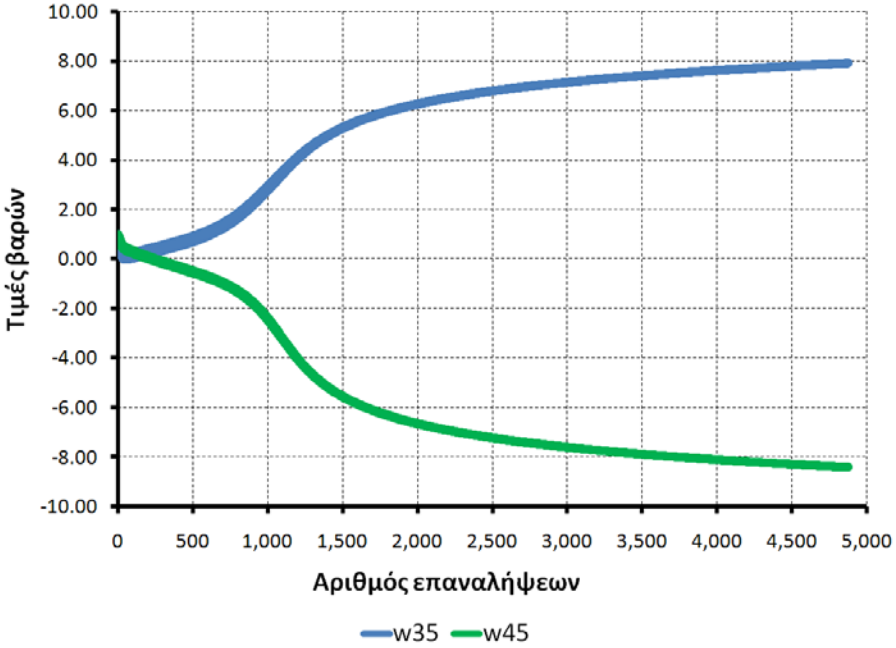
$$w_{13,new} = w_{13} + \delta_3 \cdot x_1 = +1.00$$
$$w_{23,new} = w_{23} + \delta_3 \cdot x_2 = +1.00$$
$$w_{14,new} = w_{14} + \delta_4 \cdot x_1 = +1.00$$
$$w_{24,new} = w_{24} + \delta_4 \cdot x_2 = +1.00$$
$$w_{35,new} = w_{35} + \delta_5 \cdot y_3 = +0.94$$
$$w_{45,new} = w_{45} + \delta_5 \cdot y_4 = +0.98$$
$$w_{03,new} = w_{03} + \delta_3 = +0.98$$
$$w_{04,new} = w_{04} + \delta_4 = -1.02$$
$$w_{05,new} = w_{05} + \delta_5 = +0.92$$

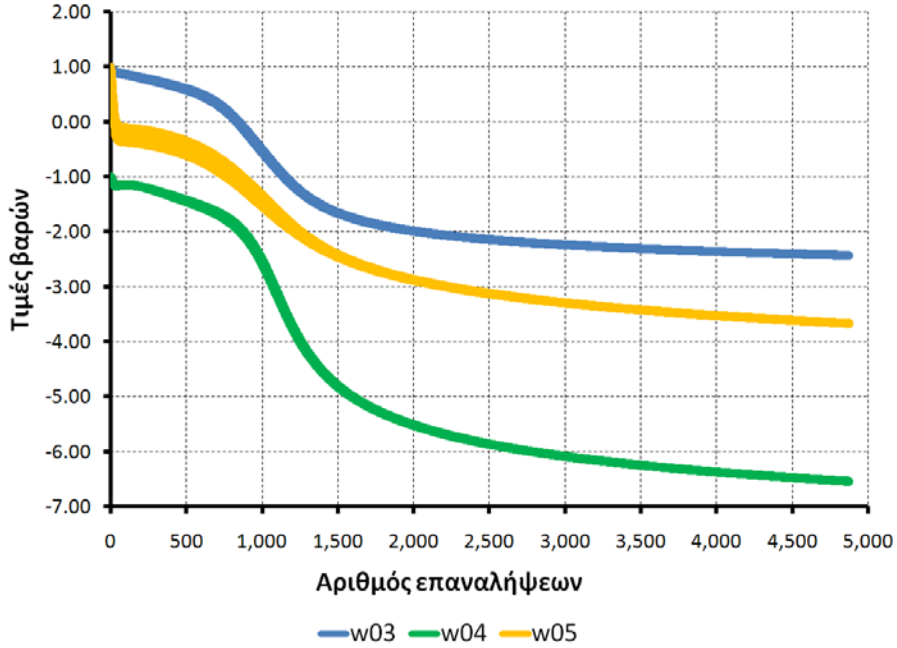


	Δεδομένα																						new								
	x1	x2	t	w13	w14	w23	w24	w03	w04	w35	w45	w05	Σ3	Σ4	y3	y4	Σ5	y5	δ5	δ4	δ3	w13'	w14'	w23'	w24'	w03'	x04'	w35'	w45'	w05'
1	0	0	0	1.00	1.00	1.00	1.00	1.00	-1.00	1.00	1.00	1.00	1.00	-1.00	0.73	0.27	2.00	0.88	-0.09	-0.02	-0.02	1.00	1.00	1.00	1.00	0.98	-1.02	0.93	0.98	0.91
2	1	0	1	1.00	1.00	1.00	1.00	0.98	-1.02	0.93	0.98	0.91	1.98	-0.02	0.88	0.50	2.21	0.90	0.01	0.00	0.00	1.00	1.00	1.00	1.00	0.98	-1.02	0.94	0.98	0.92
3	0	1	1	1.00	1.00	1.00	1.00	0.98	-1.02	0.94	0.98	0.92	1.98	-0.02	0.88	0.50	2.23	0.90	0.01	0.00	0.00	1.00	1.00	1.00	1.00	0.98	-1.01	0.95	0.98	0.92
4	1	1	0	1.00	1.00	1.00	1.00	0.98	-1.01	0.95	0.98	0.92	2.99	0.99	0.95	0.73	2.54	0.93	-0.06	-0.01	0.00	1.00	0.99	1.00	0.99	0.98	-1.03	0.89	0.94	0.86
5	0	0	0	1.00	0.99	1.00	0.99	0.98	-1.03	0.89	0.94	0.86	0.98	-1.03	0.73	0.26	1.76	0.85	-0.11	-0.02	-0.02	1.00	0.99	1.00	0.99	0.96	-1.05	0.81	0.91	0.76
6	1	0	1	1.00	0.99	1.00	0.99	0.96	-1.05	0.81	0.91	0.76	1.96	-0.06	0.88	0.49	1.91	0.87	0.01	0.00	0.00	1.00	0.99	1.00	0.99	0.96	-1.04	0.82	0.92	0.77
7	0	1	1	1.00	0.99	1.00	0.99	0.96	-1.04	0.82	0.92	0.77	1.96	-0.05	0.88	0.49	1.94	0.87	0.01	0.00	0.00	1.00	0.99	1.00	0.99	0.96	-1.04	0.84	0.92	0.78
8	1	1	0	1.00	0.99	1.00	0.99	0.96	-1.04	0.84	0.92	0.78	2.96	0.95	0.95	0.72	2.24	0.90	-0.08	-0.01	0.00	1.00	0.98	1.00	0.98	0.96	-1.05	0.76	0.87	0.71
9	0	0	0	1	0.98	1	0.98	0.96	-1.05	0.76	0.87	0.71	0.96	-1.05	0.72	0.26	1.48	0.81	-0.12	-0.02	-0.02	1	0.98	1	0.98	0.94	-1.07	0.67	0.84	0.58
10	1	0	1	1	0.98	1	0.98	0.94	-1.07	0.67	0.84	0.58	1.94	-0.1	0.87	0.48	1.57	0.83	0.02	0.01	0.00	1	0.98	1	0.98	0.94	-1.07	0.69	0.85	0.61
11	0	1	1	1	0.98	1	0.98	0.94	-1.07	0.69	0.85	0.61	1.94	-0.09	0.87	0.48	1.62	0.83	0.02	0.00	0.00	1	0.98	1	0.98	0.95	-1.06	0.71	0.86	0.63
12	1	1	0	1	0.98	1	0.98	0.95	-1.06	0.71	0.86	0.63	2.94	0.9	0.95	0.71	1.92	0.87	-0.10	-0.02	0.00	0.99	0.97	0.99	0.97	0.94	-1.08	0.62	0.79	0.53
4869	0	0	0	5.66	4.25	5.65	4.26	-2.43	-6.54	7.91	-8.41	-3.67	-2.43	-6.54	0.08	0.00	-3.04	0.05	0.00	0.00	0.00	5.66	4.25	5.65	4.26	-2.43	-6.54	7.91	-8.41	-3.67
4870	1	0	1	5.66	4.25	5.65	4.26	-2.43	-6.54	7.91	-8.41	-3.67	3.23	-2.28	0.96	0.09	3.17	0.96	0.00	0.00	0.00	5.66	4.25	5.65	4.26	-2.43	-6.54	7.92	-8.41	-3.67
4871	0	1	1	5.66	4.25	5.65	4.26	-2.43	-6.54	7.92	-8.41	-3.67	3.23	-2.28	0.96	0.09	3.17	0.96	0.00	0.00	0.00	5.66	4.25	5.65	4.26	-2.43	-6.54	7.92	-8.41	-3.66
4872	1	1	0	5.66	4.25	5.65	4.26	-2.43	-6.54	7.92	-8.41	-3.66	8.88	1.97	1.00	0.88	-3.12	0.04	0.00	0.00	0.00	5.66	4.25	5.65	4.26	-2.43	-6.54	7.92	-8.41	-3.67
	Μέσος όρος (w)			5.66	4.25	5.65	4.26	-2.43	-6.54	7.92	-8.41	-3.67																		


x1
x2
3
4
5
y
w13
w23
w14
w24
w03
w04
w05
w35
w45

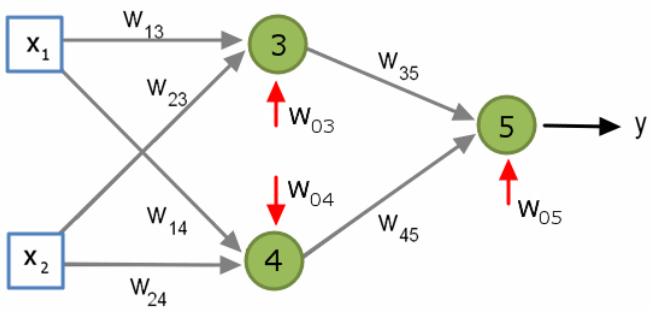

Τιμές βαρών
Αριθμός επαναλήψεων
w13
w14
w23
w24

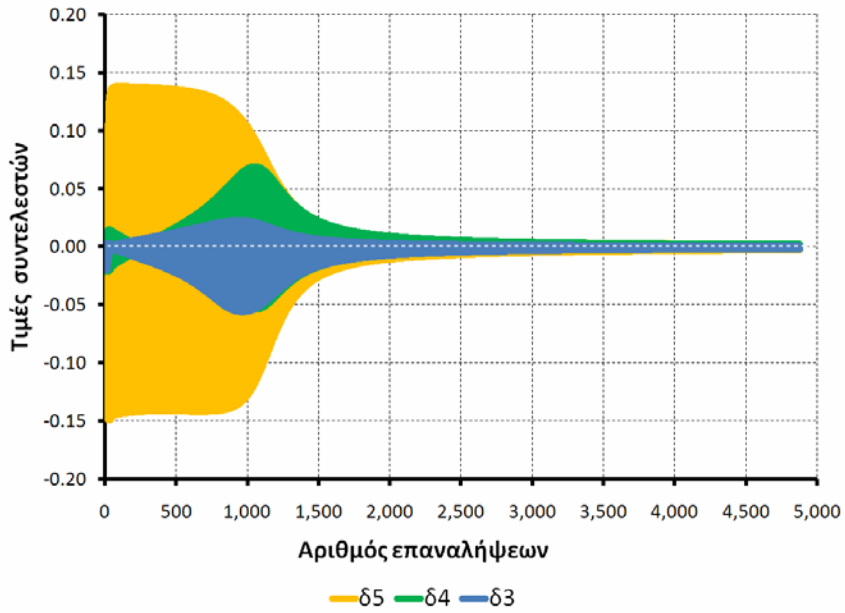

Τιμές βαρών
Αριθμός επαναλήψεων
w35
w45


Τιμές βαρών
Αριθμός επαναλήψεων
w03
w04
w05

Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	0

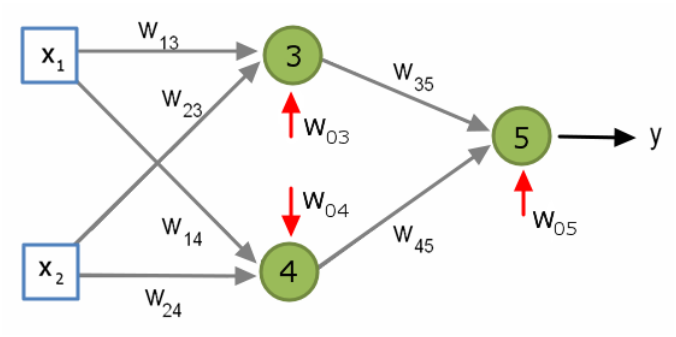




$$w_{13} = +5.66$$
$$w_{23} = +5.65$$
$$w_{14} = +4.25$$
$$w_{24} = +4.26$$
$$w_{35} = +7.92$$
$$w_{45} = -8.41$$
$$w_{03} = -2.43$$
$$w_{04} = -6.54$$
$$w_{05} = -3.67$$

Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	0



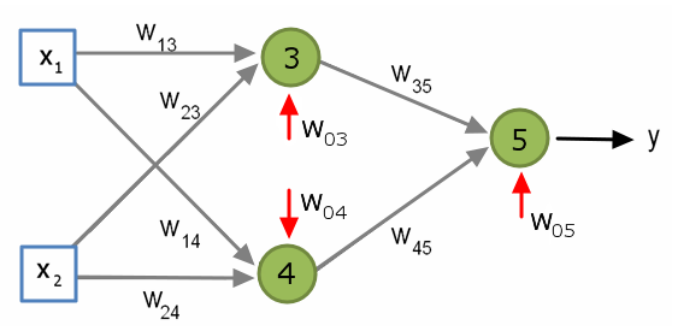
Αποτελέσματα

X1	X2	Y
0	0	0.05
1	0	0.96
0	1	0.96
1	1	0.04

$$w_{13} = +5.66$$
$$w_{23} = +5.65$$
$$w_{14} = +4.25$$
$$w_{24} = +4.26$$
$$w_{35} = +7.92$$
$$w_{45} = -8.41$$
$$w_{03} = -2.43$$
$$w_{04} = -6.54$$
$$w_{05} = -3.67$$

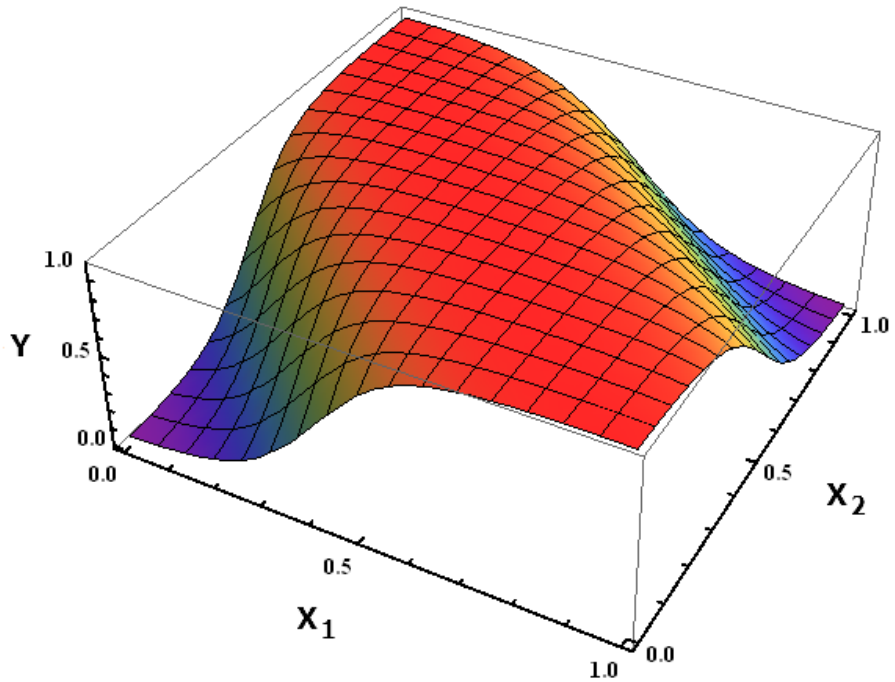
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)

X1	X2	Y
0	0	0
1	0	1
0	1	1
1	1	0

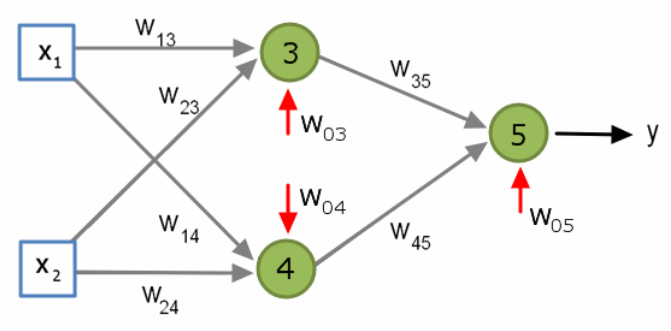


Αποτελέσματα

X1	X2	Y
0	0	0.05
1	0	0.96
0	1	0.96
1	1	0.04

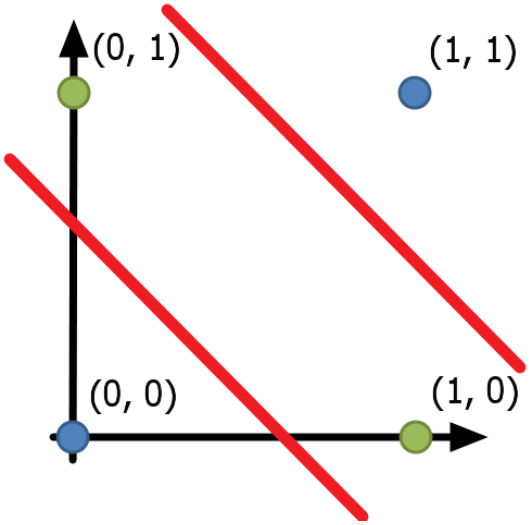


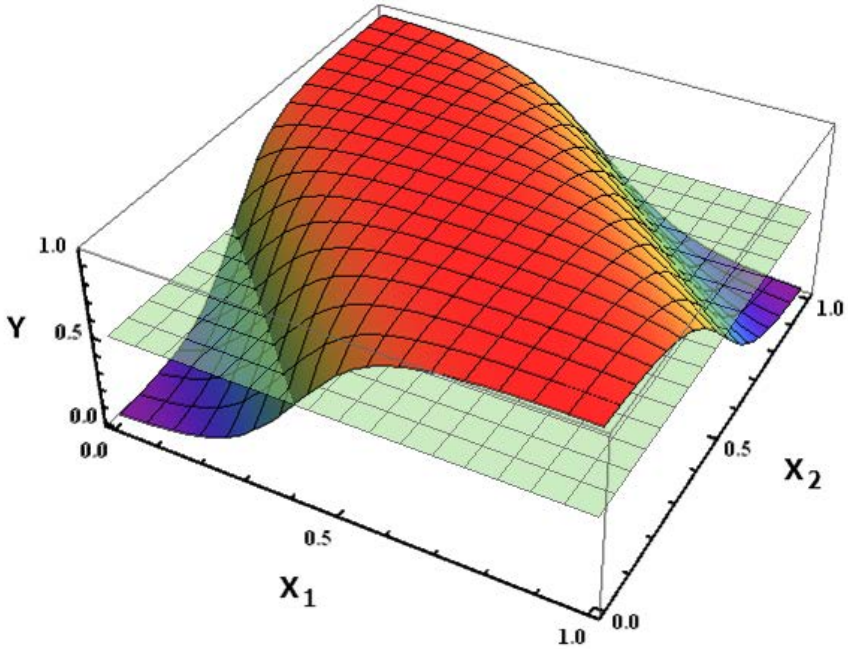
Παράδειγμα 2 (ΤΝΔ δύο επιπέδων)



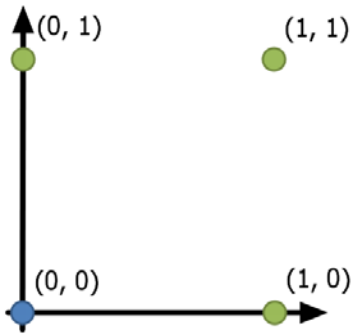
Αποτελέσματα

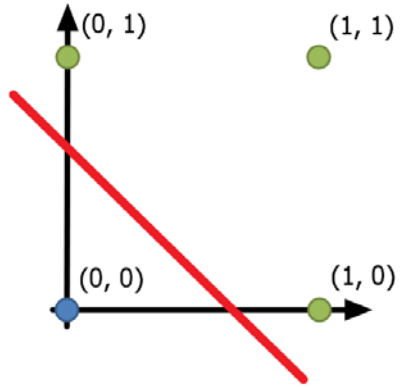
X1	X2	Y
0	0	0.05
1	0	0.96
0	1	0.96
1	1	0.04

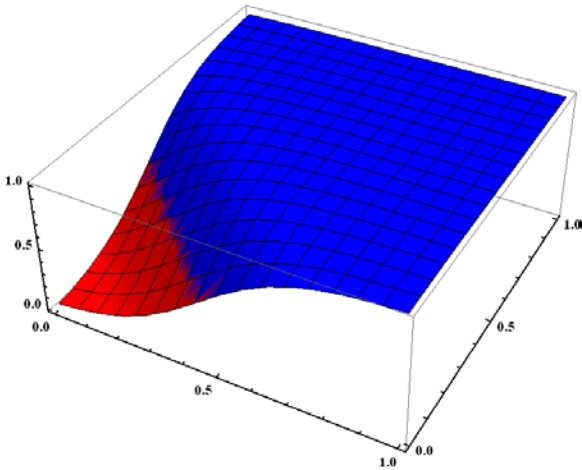


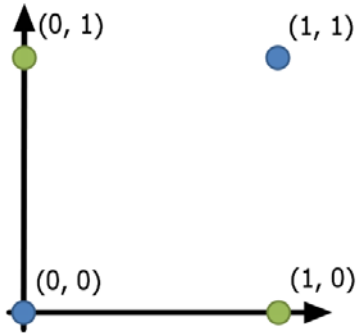


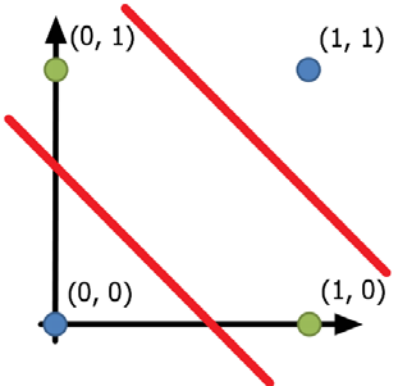
Παραδείγματα 1 & 2

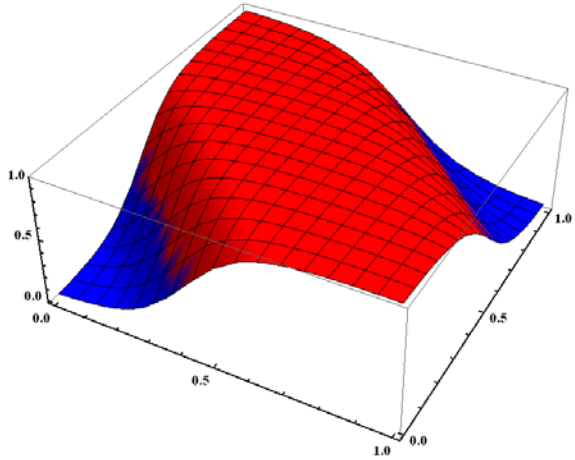




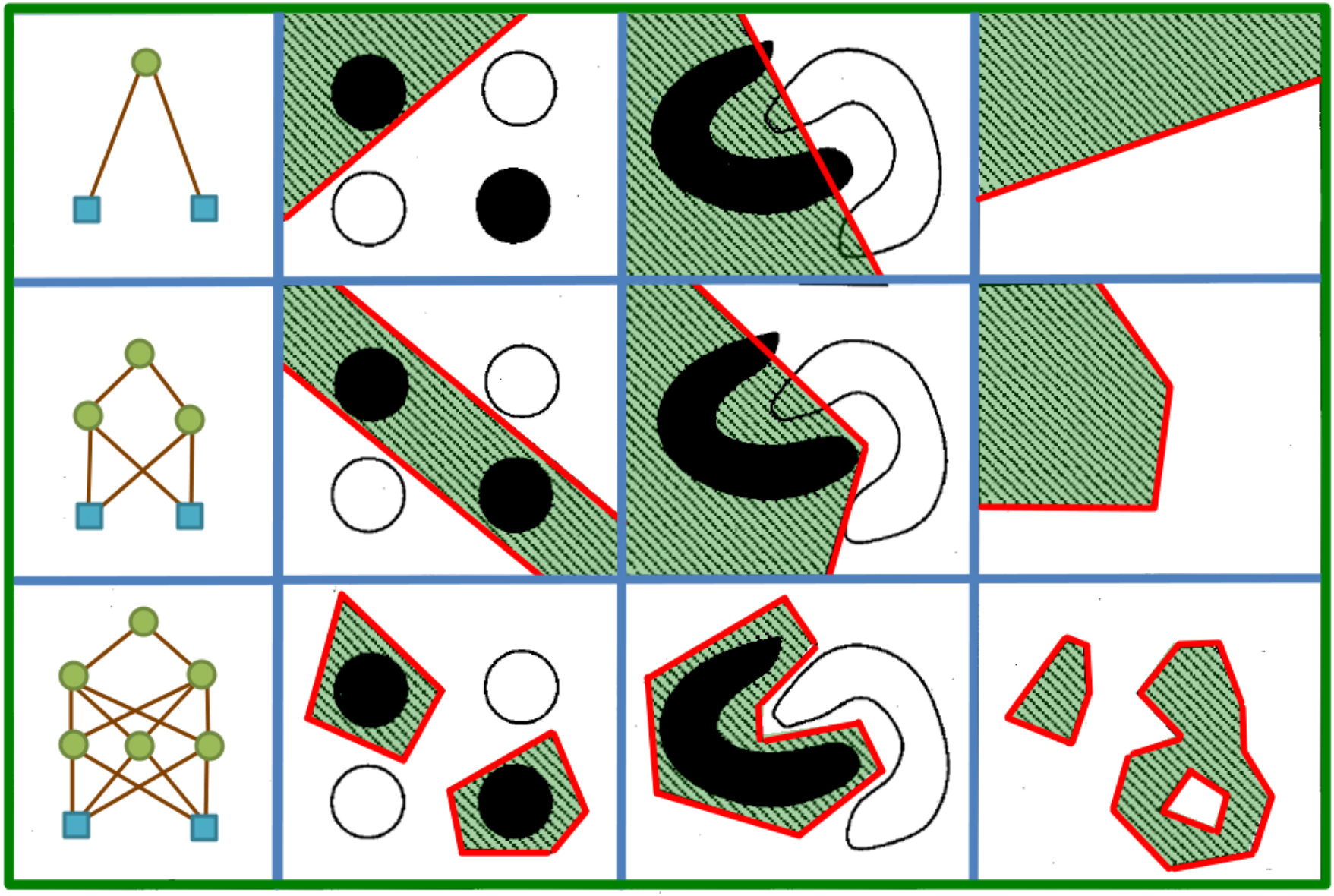








Νευρωνικά Δίκτυα



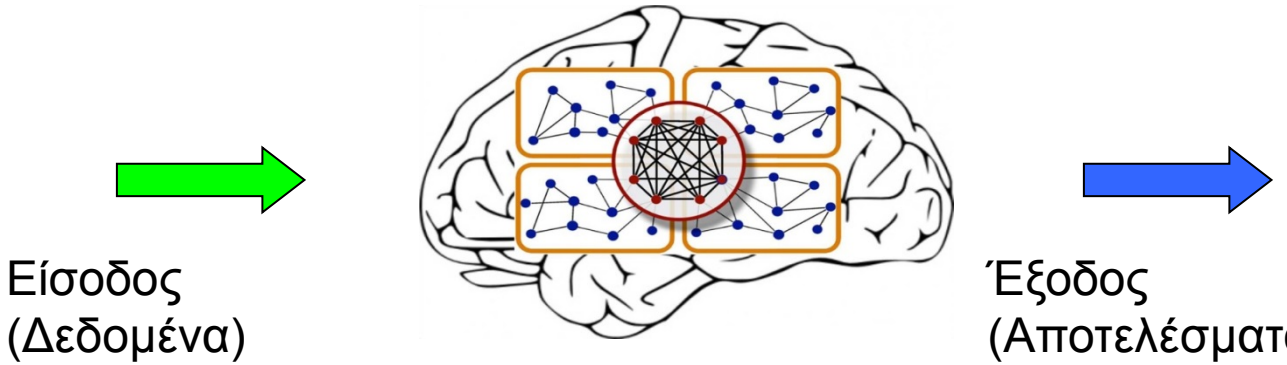
Τεχνητά Νευρωνικά Δίκτυα - Βελτιστοποίηση

Τα νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν και για την επίλυση προβλημάτων βελτιστοποίησης (π.χ., το πρόβλημα του πλανόδιου πωλητή)

- εκπαιδεύοντας ένα νευρωνικό δίκτυο με επιλυμένα προβλήματα,
- χρησιμοποιώντας δίκτυο με ανατροφοδότηση.

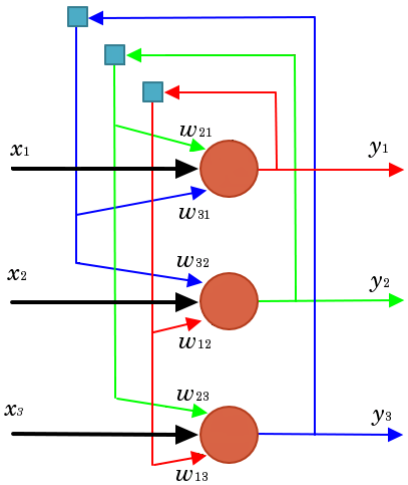
Είναι γεγονός όμως ότι τα νευρωνικά δίκτυα

- δεν αποτελούν τον αποτελεσματικότερο τρόπο επίλυσης προβλημάτων βελτιστοποίησης,
- ωστόσο όμως παρέχουν ένα εναλλακτικό τρόπο επίλυσης αυτών των προβλημάτων.



Τεχνητά Νευρωνικά Δίκτυα - Βελτιστοποίηση

Νευρωνικό δίκτυο με ανατροφοδότηση (Hopfield)



$$y_i(t+1) = \begin{cases} 1 & \alpha\nu \quad \sum_{j=1}^{N} w_{ij} \cdot x_j(t) + \theta_i > 0 \\ 0 & \alpha\nu \quad \sum_{j=1}^{N} w_{ij} \cdot x_j(t) + \theta_i < 0 \end{cases}$$

$$E(y) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{i=1}^{N} w_{ij} \cdot x_i \cdot x_j - \sum_{i=1}^{N} \theta_i \cdot x_i$$

Το δίκτυο Hopfield χαρακτηρίζεται από μία παράμετρο E, που εξαρτάται από τα σήματα εισόδου και τα βάρη, και η οποία είναι αντίστοιχη της ενέργειας σε φυσικά προβλήματα

w_{ij}	Το συνοπτικό βάρος που συνδέει τον νευρώνα i με τον νευρώνα j.
x_i	Η τιμή εισόδου του νευρώνα i.
y_i	Η τιμή εξόδου του νευρώνα i.
θ_i	Η τιμή κατωφλίου
E	Η συνάρτηση ενεργείας του νευρωνικού δικτύου
t	Η επανάληψη t

Τεχνητά Νευρωνικά Δίκτυα - Βελτιστοποίηση

Το πρόβλημα της ανάθεσης εργασιών

S1=0,5+4+5+5,5+5+4,5 = 24,5

	1	2	3	4	5	6
A	0,5	0,5	1,0	2,5	3,5	1,5
B	5,0	4,0	5,0	3,5	4,0	3,0
C	0,5	0,5	5,0	8,0	1,0	5,0
D	1,0	7,5	8,5	5,5	7,0	6,5
E	5,5	8,0	3,0	5,5	5,0	2,0
F	2,0	7,0	2,0	1,5	2,0	4,5

	1	2	3	4	5	6
A	0,5	0,5	1,0	2,5	3,5	1,5
B	5,0	4,0	5,0	3,5	4,0	3,0
C	0,5	0,5	5,0	8,0	1,0	5,0
D	1,0	7,5	8,5	5,5	7,0	6,5
E	5,5	8,0	3,0	5,5	5,0	2,0
F	2,0	7,0	2,0	1,5	2,0	4,5

Συμφώνα με το πρόβλημα, υπάρχει ίδιος αριθμός εργασιών (αριθμοί 1, 2, 3, 4, 5 και 6) και εργαζομένων (γράμματα A, B, C, D, E και F), κάθε εργασία μπορεί να εκτελεστεί από όλους τους εργαζόμενους, σε διαφορετικό χρόνο.

Σκοπός του προβλήματος είναι η ανάθεση μιας μονό εργασίας σε κάθε εργαζόμενο, με τέτοιο τρόπο ώστε να γίνεται η καλύτερη αξιοποίηση των εργαζομένων για την πραγματοποίηση όλων των εργασιών.

Τεχνητά Νευρωνικά Δίκτυα - Βελτιστοποίηση

Το πρόβλημα της ανάθεσης εργασιών

Πίνακας Κόστους Cij

	1	2	3	4	5	6
A	0.5	0.5	1	2.5	3.5	1.5
B	5	4	5	3.5	4	3
C	0.5	0.5	5	8	1	5
D	1	7.5	8.5	5.5	7	6.5
E	5.5	8	3	5.5	5	2
F	2	7	2	1.5	2	4.5

Απόκριση Νευρώνων Xij

0	1	1	0	0	1
0	0	0	1	0	0
0	0	0	0	1	0
1	0	0	0	0	1
0	0	0	1	1	1
0	0	1	0	0	0

(Cij)x(Xij)

0	0.5	1	0	0	1.5
0	0	0	3.5	0	0
0	0	0	0	1	0
1	0	0	0	0	6.5
0	0	0	5.5	5	2
0	0	2	0	0	0

$$E(y) = A \cdot \sum_{i=1}^{n} \sum_{i=1}^{n} c_{ij} \cdot x_{ij} + B \cdot \sum_{i=1}^{n} \left(\sum_{j=1}^{n} x_{ij} - 1 \right) + C \cdot \sum_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij} - 1 \right)$$

i	Οι εργαζόμενοι (γραμμή πινάκα)
j	Οι εργασίες (στήλη πινάκα)
c_{ij}	Ο πινάκας κόστος (ο χρόνος εκτέλεσης κάθε εργασίας από τον κάθε εργαζόμενο)
x_{ij}	Η τιμή εισόδου νευρώνα (για τον εργαζόμενο i, στην εργασία j)
n	Ο αριθμός εργασιών και εργαζομένων (διάσταση πινάκα n=6)
A, B, C	Όροι ποινής, (για τη συνάρτηση ενέργειας), A=1, B=C=4

Τεχνητά Νευρωνικά Δίκτυα - Βελτιστοποίηση

t=0

Πίνακας Κόστους Cij

	1	2	3	4	5	6
A	0.5	0.5	1	2.5	3.5	1.5
B	5	4	5	3.5	4	3
C	0.5	0.5	5	8	1	5
D	1	7.5	8.5	5.5	7	6.5
E	5.5	8	3	5.5	5	2
F	2	7	2	1.5	2	4.5

Απόκριση Νευρώνων Xij

0	1	1	0	0	1	2
0	0	0	1	0	0	0
0	0	0	0	1	0	0
1	0	0	0	0	1	1
0	0	0	1	1	1	2
0	0	1	0	0	0	0

(Άθροισμα Στηλών) -1

S3=5

(Άθροισμα Γραμμών) - 1

0	0	1	1	1	2	S2=5

(Cij)x(Xij)

0	0.5	1	0	0	1.5
0	0	0	3.5	0	0
0	0	0	0	1	0
1	0	0	0	0	6.5
0	0	0	5.5	5	2
0	0	2	0	0	0

Άθροισμα Κελιών Πίνακα

S1 = 29.5

E = 1xS1 + 4xS2 + 4xS3 = 69.5

t=1

Πίνακας Κόστους Cij

	1	2	3	4	5	6
A	0.5	0.5	1	2.5	3.5	1.5
B	5	4	5	3.5	4	3
C	0.5	0.5	5	8	1	5
D	1	7.5	8.5	5.5	7	6.5
E	5.5	8	3	5.5	5	2
F	2	7	2	1.5	2	4.5

Απόκριση Νευρώνων Xij

0	1	1	0	0	1	2
0	0	0	1	0	0	0
0	0	1	0	1	0	1
1	0	0	0	0	1	1
0	0	0	1	1	1	2
0	0	1	0	0	0	0

(Άθροισμα Στηλών) -1

S3=6

(Άθροισμα Γραμμών) - 1

0	0	2	1	1	2	S2=6

(Cij)x(Xij)

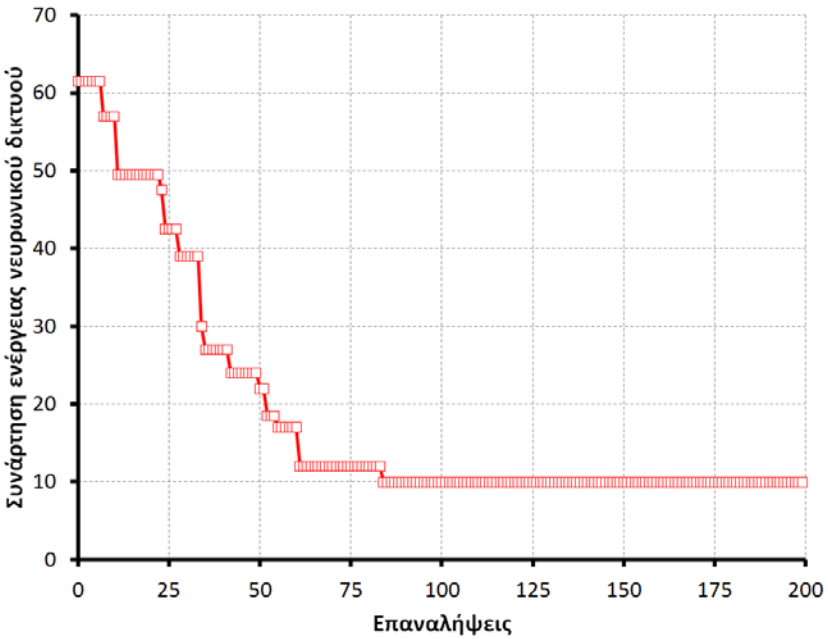
0	0.5	1	0	0	1.5
0	0	0	3.5	0	0
0	0	5	0	1	0
1	0	0	0	0	6.5
0	0	0	5.5	5	2
0	0	2	0	0	0

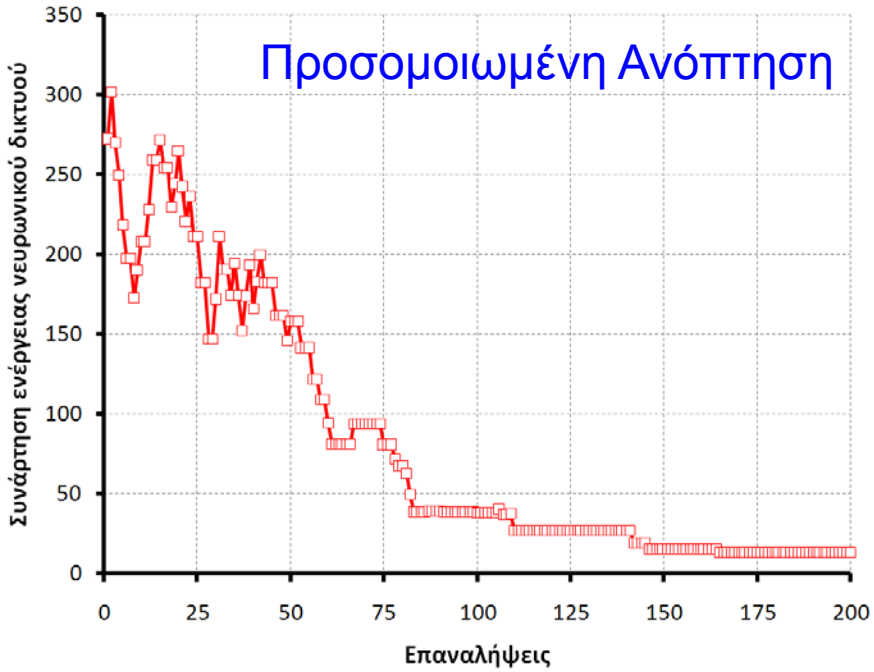
Άθροισμα Κελιών Πίνακα

S1 = 34.5

E = 1xS1 + 4xS2 + 4xS3 = 82.5

Τεχνητά Νευρωνικά Δίκτυα - Βελτιστοποίηση





Πίνακας Κόστους Cij

	1	2	3	4	5	6
A	0.5	0.5	1	2.5	3.5	1.5
B	5	4	5	3.5	4	3
C	0.5	0.5	5	8	1	5
D	1	7.5	8.5	5.5	7	6.5
E	5.5	8	3	5.5	5	2
F	2	7	2	1.5	2	4.5

Απόκριση Νευρώνων Xij

0	1	0	0	0	0
0	0	0	0	0	1
0	0	0	0	1	0
1	0	0	0	0	0
0	0	1	0	0	0
0	0	0	1	0	0

0	(Άθροισμα Στηλών) -1
0	
0	
0	
0	
0	
S3=0	

(Cij)x(Xij)

0	0.5	0	0	0	0
0	0	0	0	0	3
0	0	0	0	1	0
1	0	0	0	0	0
0	0	3	0	0	0
0	0	0	1.5	0	0

(Άθροισμα Γραμμών) - 1

0	0	0	0	0	0	S2=0

Άθροισμα Κελιών Πίνακα

S1 = 10

βέλτιστη λύση

E = 1xS1 + 4xS2 + 4xS3 = 10

Τεχνητά Νευρωνικά Δίκτυα

- Τα νευρωνικά δίκτυα προσπαθούν να συνδυάσουν τον τρόπο σκέψης του ανθρώπινου εγκεφάλου με τον αφηρημένο μαθηματικό τρόπο σκέψης.
- Ένα νευρωνικό δίκτυο είναι ένα σύστημα το οποίο επεξεργάζεται και εξάγει αποτελέσματα ενώ παράλληλα μπορεί
  - να αποθηκεύει γνώση,
  - να εφαρμόζει (χρησιμοποιεί) την αποθηκευμένη γνώση,
  - να αποκτά νέα γνώση μέσω εμπειρίας.
- Ένα νευρωνικό δίκτυο παρουσιάζει ανοχή σε λάθη επειδή η επεξεργασία διαχέεται σε όλη την έκταση των συνδέσεων του δικτύου – αντίθετα, ο προγραμματισμός που βασίζεται σε μια ακολουθία συνθηκών (if-then) μπορεί να έχει προβλήματα σε περίπτωση λαθών στα δεδομένα.
- Η ευελιξία των νευρωνικών δικτύων αποτελεί ταυτόχρονα και αδυναμία επειδή
  - όσο μεγαλύτερο είναι το δίκτυο, τόσο πιο «εύθραυστο» είναι,
  - όσο μεγαλύτερο είναι, τόσο δυσκολότερα επεξεργάσιμο είναι.
- Τα νευρωνικά δίκτυα αδυνατούν να εξηγήσουν ποιοτικά τη γνώση που μοντελοποιούν.

Τέλος παρουσίασης