




Πανεπιστήμιο Αιγαίου
Πολυτεχνική Σχολή
Τμήμα Μηχανικών Οικονομίας & Διοίκησης

Εργαστήριο Διοίκησης Επιχειρήσεων &
Λήψης Αποφάσεων (ΔΕΛΑΠ)
mde-lab.aegean.gr

Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα

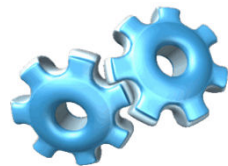
Συμπληρωματικές Διαλέξεις στο Μάθημα Ανάλυση Αποφάσεων και Μηχανική Γνώσεων



Γεώργιος Δούνιας
Καθηγητής

Ευάγγελος Καραμπότσης
Υποψήφιος Διδάκτορας

Γενικά για τα Τεχνητά Νευρωνικά Δίκτυα



Τεχνητά Νευρωνικά Δίκτυα

Νευρωνικά Δίκτυα

Συνδυαστικά Μοντέλα

Παράλληλα Μοντέλα Κατανεμημένης Επεξεργασίας

Νευρομορφικά Συστήματα



Ένα Τεχνητό Νευρωνικό Δίκτυο (ΤΝΔ) είναι ένας μεγάλος παράλληλος επεξεργαστής με κατανεμημένη αρχιτεκτονική που αποτελείται από απλές μονάδες επεξεργασίας (Τεχνητοί Νευρώνες) και έχει τη δυνατότητα να αποθηκεύει εμπειρική γνώση και να τη καθιστά διαθέσιμη για χρήση

- ✓ Τα ΤΝΔ προσομοιώνουν τη συμπεριφορά του ανθρώπινου εγκεφάλου
- ✓ Τα ΤΝΔ αποκτούν γνώση μέσω της διαδικασίας της μάθησης (Δοκιμή & Σφάλμα)
- ✓ Τα ΤΝΔ αποθηκεύουν τη γνώση στα Συνοπτικά Βάρη (Οι δυνάμεις σύνδεσης των τεχνητών νευρώνων)



Μερικά από τα Βασικά Χαρακτηριστικά των Τεχνητών Νευρωνικών Δικτύων:

- Από τις νεότερες μεθόδους επεξεργασίας σήματος πληροφορίας στο πεδίο της μηχανικής μάθησης
- Αποτελούνται από πληθώρα τεχνητών νευρώνων που λειτουργούν παράλληλα
- Συμβάλουν σε σημαντικές διαδικασίες ως i) ταξινομητές, ii) μη γραμμικά προσαρμοστικά φίλτρα, iii) προσεγγιστές συναρτήσεων, iv) ελεγκτές συστημάτων κ.λπ.
- Όταν ένας νευρώνας τούς αστοχεί κατά τη λειτουργία του, τότε τα ΤΝΔ μπορούν να συνεχίσουν να δουλεύουν, λόγω της παράλληλης φύσης τους
- Η αντίδραση του νευρώνα εξαρτάται μόνο από τοπική πληροφορία
- Ικανότητα μάθησης, ανάκλησης/γενίκευσης
- Οι συνδέσεις με τα ρυθμιζόμενα βάρη αποθηκεύουν γνώση
- Η συλλογική συμπεριφορά επιδεικνύει την υπολογιστική ισχύ



Βασικά Πλεονεκτήματα των ΤΝΔ:

- Έχουν την ικανότητα να μαθαίνουν από παραδείγματα
- Μπορούν να θεωρηθούν ως κατανεμημένη μνήμη ή μνήμη συσχέτισης
- Εξαιρετική ικανότητα αναγνώρισης προτύπων
- Μπορούν να προσομοιώσουν πολύπλοκες διεργασίες πολύ καλύτερα από άλλες γραμμικές μεθόδους.
- Μπορούν να εξάγουν πληροφορίες από τα δεδομένα, χωρίς παλαιότερη γνώση των δεδομένων ή του συστήματος που αναλύεται
- Μπορούν να μεταβάλουν τη συμπεριφορά τους ανάλογα με τις αλλαγές του συστήματος
- Μπορούν να χωρίσουν το χώρο των διαθέσιμων δεδομένων και να δημιουργήσουν διαφορετικές σχέσεις για διαφορετικά τμήματα του χώρου των δεδομένων
- Μπορούν να οργανώσουν τμηματικά μη-γραμμικά μοντέλα που να συμπεριλαμβάνουν ασυνέχειες
- Έχουν μεγάλη ανοχή στα σφάλματα των δεδομένων



Βασικά Μειονέκτημα των ΤΝΔ:

- Η επιλογή του βέλτιστου δικτύου (αρχιτεκτονική και παράμετροι) σχετίζεται με την εφαρμογή και πρέπει να γίνεται μέσω διαδικασίας trial-and-error
- Δεν είναι ικανά να ανταποκριθούν σε πολύ μεγάλες μεταβολές του συστήματος που προσπαθούν να προσομοιώσουν καθώς έχουν εκπαιδευτεί σε ιστορικά δεδομένα
- Δεν παρέχουν / επιτρέπουν την σαφή κατανόηση της απόκρισης των συστατικών ενός ΤΝΔ (αριθμός νευρώνων / στρωμάτων, βάρη) σε σύγκριση με άλλα υβριδικά μοντέλα
- Μπορεί να αποτύχουν να αναπαράγουν το επιθυμητό αποτέλεσμα λόγω:
 - ✓Αδυναμίας κατά την εκπαίδευση να προσδιοριστούν τα κατάλληλα βάρη
 - ✓Ακατάλληλου μεγέθους/αρχιτεκτονικής



Πεδία Εφαρμογών ΤΝΔ:

- **Τακτοποίηση και έλεγχο πολύπλοκων συστημάτων**
Συναρτησιακή προσέγγιση διαδικασιών
- **Φυσική υψηλών ενεργειών**
Αναγνώριση του τύπου των φυσικών σωματιδίων
- **Μετεωρολογία**
Μοντέλα πρόβλεψης της κίνησης των αερίων μαζών
- **Ρομποτική και Διαστημική τεχνολογία**
Αυτόνομα ρομποτικά συστήματα
- **Ιατρική**
Επεξεργασία και κατάτμηση ιατρικών εικόνων, Ανάλυση βιολογικών Σημάτων, συστήματα υποστήριξης ιατρικής διάγνωσης
- **Βιοπληροφορική**
Συστήματα διάγνωσης του καρκίνου με βάση εξετάσεις DNA
- **Οικονομία, Επιχειρησιακή έρευνα**
Μοντελοποίηση και πρόβλεψη χρονοσειρών σχετιζόμενων με τιμές μετοχών και οικονομικών δεικτών, συστήματα υποστήριξης αποφάσεων
- **Επεξεργασία δεδομένων πολυμέσων**
Ανάλυση, κατάτμηση και συμπίεση εικόνων και video, παρακολούθηση τροχιάς αντικειμένων, αναγνώριση γεγονότων σε video
- **Εφαρμογές Αναγνώρισης**
Αναγνώριση φωνής, αναγνώριση ομιλούντος, αναγνώριση αντικειμένων σε εικόνες και video, αναγνώριση χαρακτήρων χειρόγραφου κειμένου, αναγνώριση προσώπων
- **Διαχείριση Γνώσης**
Ταξινόμηση κειμένων με βάση το περιεχόμενο, αυτόματη ομαδοποίηση ιστοσελίδων, ανίχνευση spam



Τυπικές Εφαρμογές ΤΝΔ στα πλαίσια της Ανάλυσης Δεδομένων:

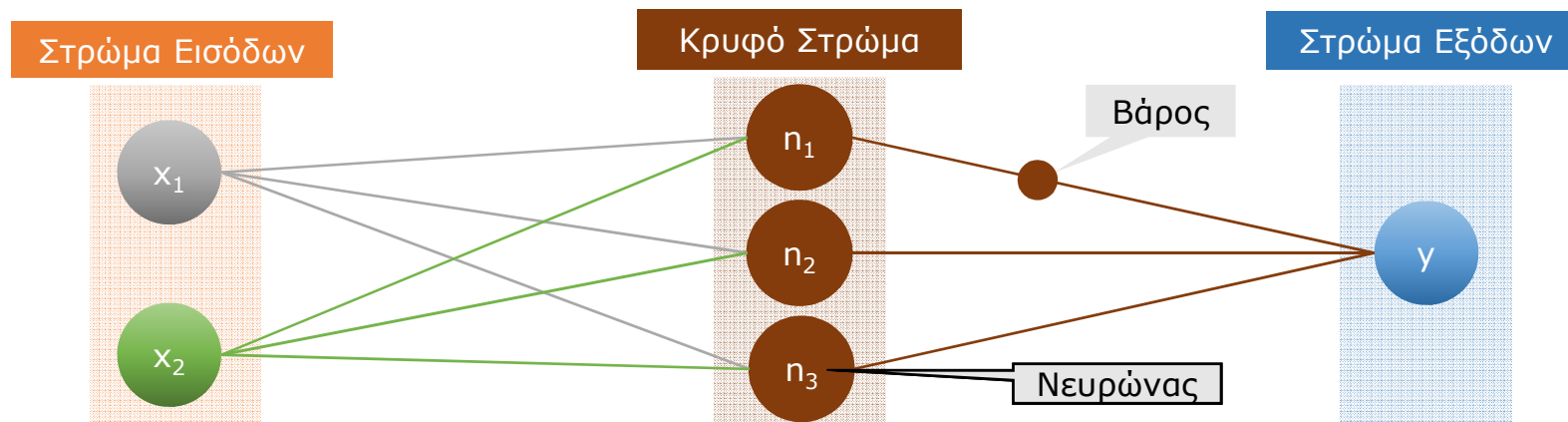
- **Ταξινόμηση**
Ιατρική διάγνωση, αναγνώριση στόχου, χαρακτήρων και φωνής
- **Προσέγγιση Συνάρτησης**
Μοντελοποίηση διαδικασίας, έλεγχος διαδικασίας, μοντελοποίηση δεδομένων, διαγνωστική βλαβών δυναμικών συστημάτων, αναγνώριση συστημάτων
- **Πρόβλεψη Χρονοσειρών**
Πρόβλεψη οικονομικών μεγεθών (πτώχευση, πωλήσεις), δυναμική μοντελοποίηση συστημάτων
- **Εξόρυξη Δεδομένων**
Ομαδοποίηση, οπτικοποίηση δεδομένων (στατιστικά γραφήματα, θεματικοί χάρτες), εκμαίευση δεδομένων



Δομικά Στοιχεία των Τεχνητών Νευρωνικών Δικτύων



- Τα ΤΝΔ είναι Μαθηματικά Μοντέλα εμπνευσμένα από τη Νευροφυσιολογία
- Τα ΤΝΔ αποτελούνται από τεχνητούς νευρώνες, δηλαδή στοιχείων επεξεργασίας της πληροφορίας
- Η αντίδραση του τεχνητού νευρώνα εξαρτάται μόνο από τη τοπική πληροφορία (δηλαδή, από τη πληροφορία που θα εισάγουμε στο ΤΝΔ)
- Τα ΤΝΔ έχουν την ικανότητα της μάθησης, της Ανάκλησης/Γενίκευσης
- Οι συνδέσεις με τα ρυθμιζόμενα βάρη αποθηκεύουν γνώση
- Τυπικό Μοντέλο ΤΝΔ ενός Κρυφού Στρώματος:



Ο Ανθρώπινος Εγκέφαλος:

- είναι το πιο πολύπλοκο όργανο του ανθρώπινου σώματος
- επιτελεί τις ανώτερες λειτουργίες, δέχεται αισθητικά ερεθίσματα και εκπέμπει κινητικές διεγέρσεις
- αποτελείται από: το Στέλεχος (Υπεύθυνο για τις ζωτικές λειτουργίες, όπως η αναπνοή, τη Παρεγκεφαλίδα (Υπεύθυνη για την ισορροπία και το συντονισμό των κινήσεων του ανθρώπινου σώματος) και από 2 εγκεφαλικά ημισφαίρια (Περιβάλλονται από το φλοιό)

Ο Εγκεφαλικός Φλοιός:

- είναι η πιο πολύπλοκη εγκεφαλική δομή
- αποτελεί την Έδρα Νόησης στον άνθρωπο
- καταλαμβάνει το μεγαλύτερο ποσοστό (85%) της συνολικής μάζας του εγκεφάλου
- είναι σημείο συγκέντρωσης των νευρώνων, σε ομάδες

- 💡 Στον εγκέφαλο, ενός νεογέννητου ανθρώπου, υπάρχουν περίπου 10^{11} νευρώνες
- 💡 Οι νευρώνες σχηματίζουν μεταξύ τους περίπου 10^{15} συνάψεις
- 💡 Ο εγκέφαλος καταναλώνει το **10%** της συνολικής ενέργειας του σώματος
- 💡 Υποστηρίζεται πως, μετά την ηλικία των **30**, καταστρέφονται περίπου **1000/ημέρα**

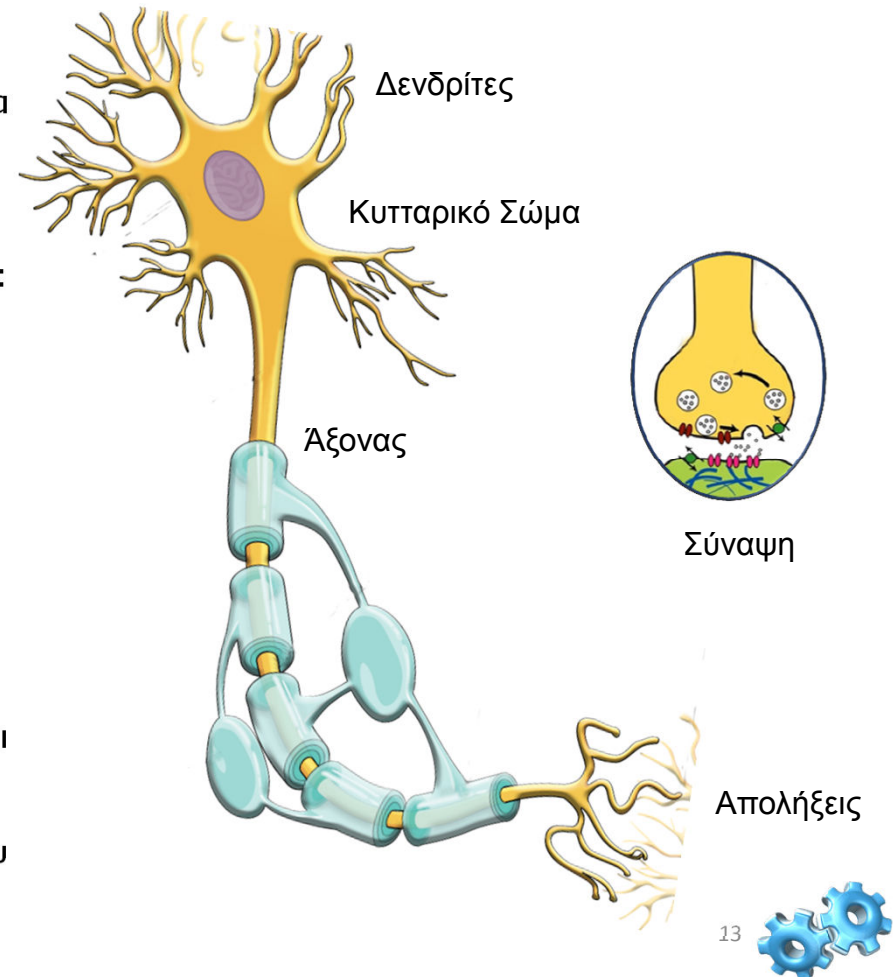


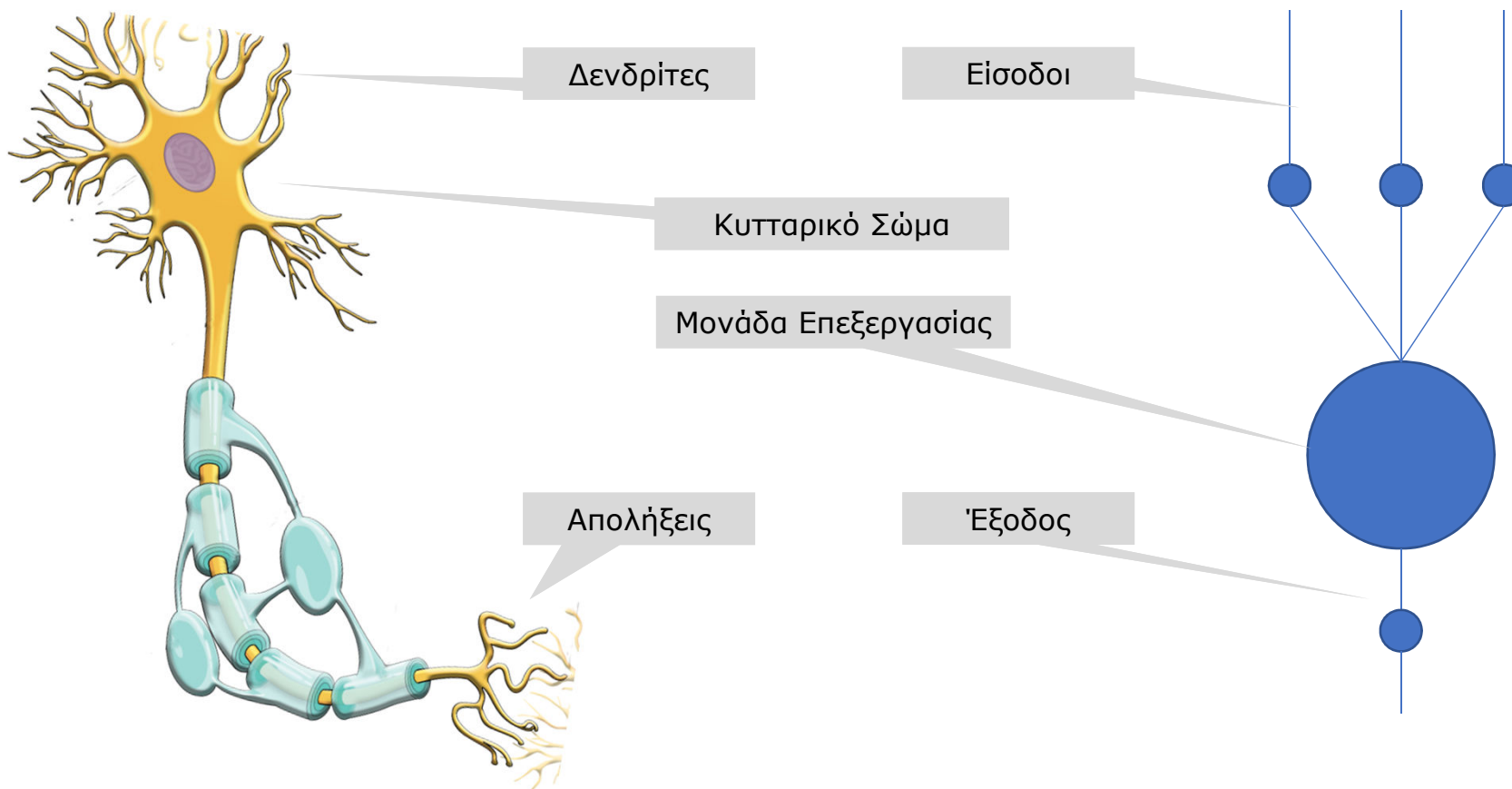
Ο Βιολογικός Νευρώνας:

- παραλαμβάνει επεξεργάζεται και εγγράφει τα ερεθίσματα στον εγκέφαλο
- μεταβιβάζει τις εντολές του στα εκτελεστικά όργανα
- αποτελείται από τα εξής Βασικά Δομικά Στοιχεία: Δενδρίτες, Κυτταρικό Σώμα, Άξονας & Αξονικές Απολήξεις

Οι Συνάψεις:

- δημιουργούνται:
 - ✓ Μεταξύ δενδριτών
 - ✓ Μεταξύ δενδριτών και κυτταρικού σώματος
 - ✓ Μεταξύ των απολήξεων δύο αξόνων
 - ✓ Μεταξύ των απολήξεων του άξονα ενός νευρώνα και του κυτταρικού σώματος ή δενδρίτη ενός άλλου
- επιτρέπουν τη δημιουργία ενός μικρού διάκενου, που ονομάζεται συνοπτικό κενό, χάσμα ή σχισμή

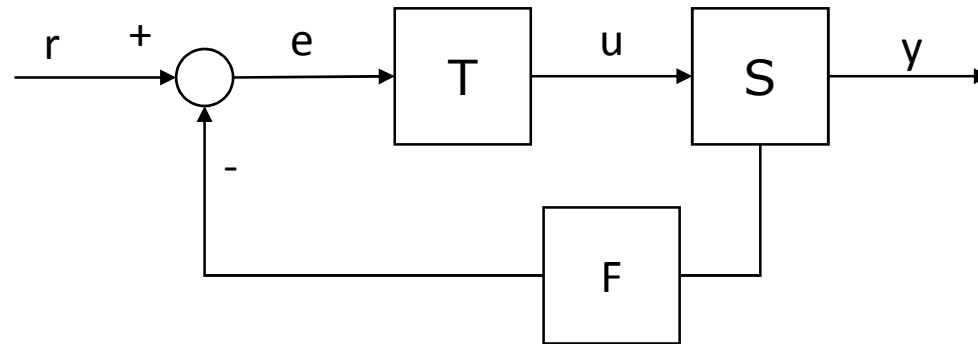




Ανάδραση & Αναπαράσταση Γνώσης στα ΤΝΔ



Ανάδραση γενικά ορίζεται η διαδικασία ανατροφοδότησης ενός σήματος εξόδου στην είσοδο ενός συγκεκριμένου συστήματος



- Η ανάδραση λαμβάνει χώρα σε όλα τα μέρη του νευρικού συστήματος κάθε ζώντος οργανισμού
- Με την ανάδραση σε ένα ΤΝΔ επιτυγχάνουμε:
 - ✓ Δημιουργία κλειστών βρόγχων
 - ✓ Διόρθωση των σφαλμάτων του



Ως Γνώση θα μπορούσε να χαρακτηριστεί η αποθηκευμένη πληροφορία που χρησιμοποιείται από έναν άνθρωπο ή μια μηχανή για να ερμηνεύσει ένα πρόβλημα, για να εκτιμήσει μια κατάσταση, να υπολογίσει μια τιμή, γενικά για να αντιδράσει σε ερεθίσματα του εξωτερικού του κόσμου.

Ένα ΤΝΔ πρέπει να κατασκευάσει ένα μοντέλο του περιβάλλοντος στο οποίο λειτουργεί και να φροντίσει για τη διατήρηση της συνέπειας του μοντέλου με το περιβάλλον. Η γνώση του περιβάλλοντος ενός ΤΝΔ αποτελείται από δυο (2) είδη πληροφορίας:

- Πρωταρχική πληροφορία για τη κατάσταση του περιβάλλοντος
- Πληροφορία που προέρχεται από παρατηρήσεις (μετρήσεις, συνήθως θορυβώδεις) προερχόμενες από το περιβάλλον που παρέχουν πληροφορίες για το περιβάλλον και χρησιμοποιούνται ως παραδείγματα/πρότυπα εκπαίδευση ενός ΤΝΔ. Τα εν λόγω παραδείγματα είναι δεδομένα εισόδων και επιθυμητών εξόδων ενός ΤΝΔ.



Η διαδικασία της αναπαράστασης της γνώσης σε ένα ΤΝΔ χωρίζεται σε δύο (2) φάσεις:

- Φάση της Εκπαίδευσής (Learning)
- Φάση της Ανάκλησης/Γενίκευσης (Recall/Generalization)

Σε ένα ΤΝΔ συγκεκριμένης αρχιτεκτονικής η γνώση του περιβάλλοντος κωδικοποιείται στις παραμέτρους του δικτύου (συνοπτικά βάρη και σταθερές πόλωσης).



Μοντέλα Τεχνητών Νευρώνων



Γενικά

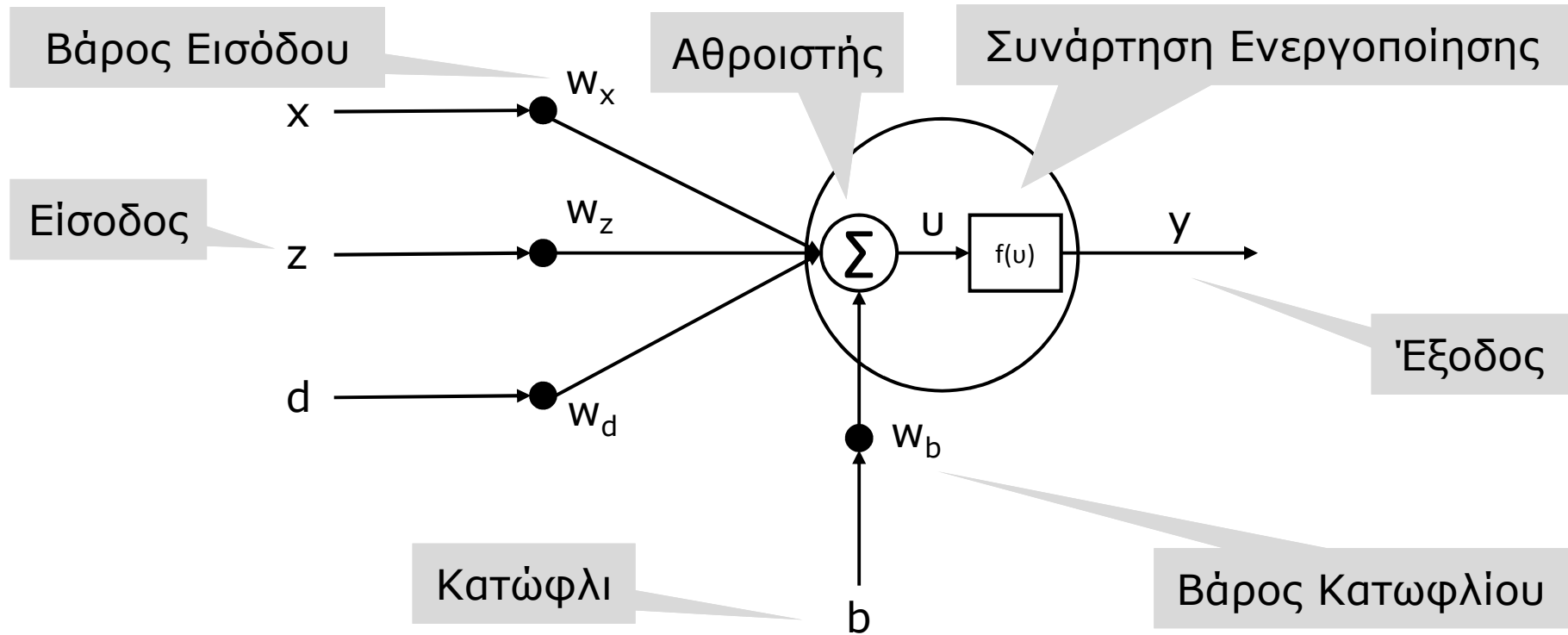
MP Neuron

Perceptron

ADALINE

Στατικός

Δυναμικός



Το Μοντέλο του Τεχνητού Νευρώνα:

- δέχεται **Σήματα Εισόδου** (x_n) με αντίστοιχα **Βάρη** (w_n)
 - ✓ Τα σήματα εισόδου αντιστοιχούν σε συνεχείς μεταβλητές, σε αντίθεση με τους ηλεκτρικούς παλμούς του εγκεφάλου
 - ✓ Τα βάρη μεταβάλουν τα σήματα εισόδου και μπορεί να είναι αρνητικά ή θετικά σε αντιστοιχία της επιβραδυντικής ή επιταχυντικής λειτουργίας των συνάψεων ενός βιολογικού νευρώνα
- έχει το **Κατώφλι/Πόλωση** (b)
- αποτελείται από τον **Αθροιστή** (Σ) και τη **Συνάρτηση Ενεργοποίησης** $f(u)$
 - ✓ Ο αθροιστής αθροίζει τα (επηρεασμένα από τα βάρη) σήματα εισόδου ($u = \sum_{i=0,1,\dots,n} w_i x_i$)
 - ✓ Η συνάρτηση ενεργοποίησης διαμορφώνει τη τελική τιμή του σήματος εξόδου ($y = f(u)$)
- παράγει ένα **Σήμα Εξόδου** (y)
 - ✓ Η τελική τιμή εξόδου είναι μοναδική



Πόλωση Τεχνητού Νευρώνα:

- Μεροληψία (Bias) ή Κατώφλι (Threshold)
- Το κατώφλι ορίζεται ως σήμα εισόδου (Συνήθως $b=0$, $b=1$ ή -1)
- Η τιμή του βάρους κατωφλίου (w_b) καθορίζει έμμεσα και δυναμικά τη θέση της συνάρτησης ενεργοποίησης στο καρτεσιανό σύστημα
- Ελάττωση του b έχει ως αποτέλεσμα μετατόπιση της εξόδου y δεξιά
- Αύξηση του w_b έχει ως αποτέλεσμα τη συρρίκνωση της εξόδου y

Τυπικές Συναρτήσεις Ενεργοποίησης:

- Βηματική Συνάρτηση:

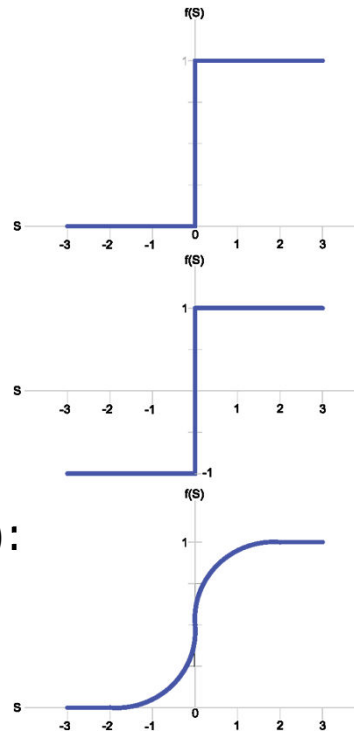
$$f(v) = \begin{cases} 1, & \text{αν } v > b \\ 0, & \text{αν } v \leq b \end{cases}$$

- Συνάρτηση Προσήμου:

$$f(v) = \begin{cases} 1, & \text{αν } v > b \\ -1, & \text{αν } v \leq b \end{cases}$$

- Λογιστική Συνάρτηση (Σιγμοειδής):

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



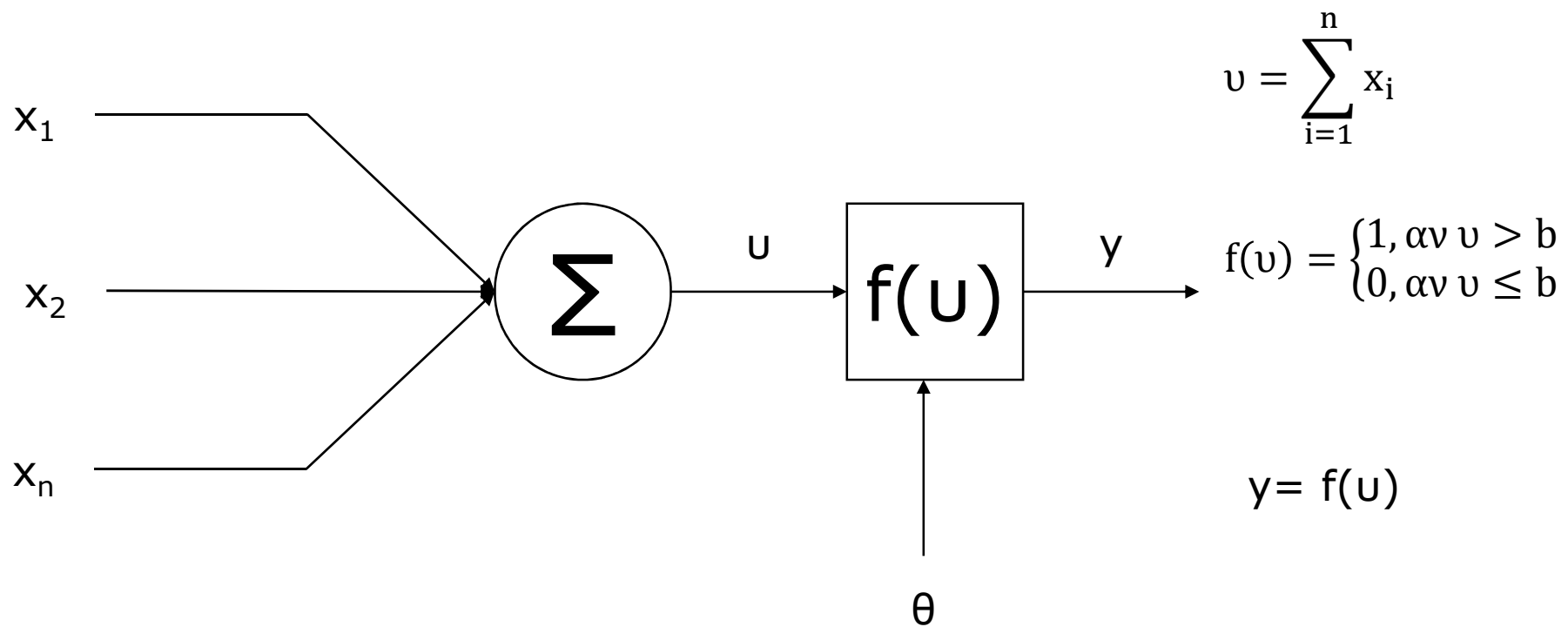
💡 Οι νευρώνες χαρακτηρίζονται σύμφωνα με τη συνάρτηση ενεργοποίησης σε γραμμικούς, μη γραμμικούς και στοχαστικούς

💡 Οι παραπάνω συναρτήσεις είναι μη γραμμικές

Νευρώνας McCulloch and Pitts

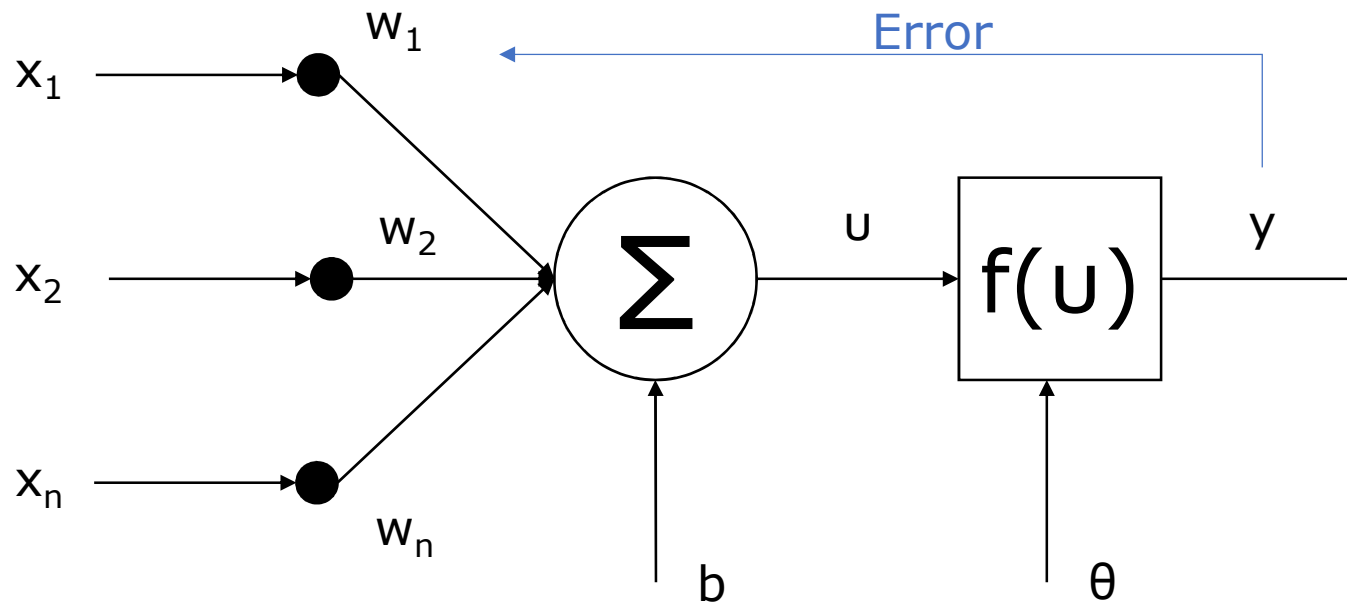
- Προτάθηκε από τους McCulloch and Pitts το 1943
- Μιμείται τη λειτουργία ενός βιολογικού νευρώνα
- Δέχεται δυαδικές εισόδους $\{0,1\}$
- Παράγει δυαδικές εξόδους $\{0,1\}$ βάση μιας τιμής κατωφλίου θ που ρυθμίζεται
- Δεν έχει συνοπτικά βάρη
- Έχει ως συνάρτηση ενεργοποίησης τη βηματική συνάρτηση
- Μπορεί να χρησιμοποιηθεί για προβλήματα ταξινόμησης





Νευρώνας Perceptron

- Προτάθηκε από τον Frank Rosenblatt το 1957
- Είναι μια ελαφρώς τροποποιημένη έκδοση του νευρώνα MP
- Δέχεται πραγματικές τιμές ως εισόδους
- Παράγει δυαδικές εξόδους $\{0,1\}$ βάση μιας τιμής κατωφλίου θ , που ρυθμίζεται
- Έχει συνοπτικά βάρη
- Έχει ως συνάρτηση ενεργοποίησης τη βηματική συνάρτηση
- Μπορεί να χρησιμοποιηθεί για προβλήματα ταξινόμησης



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

$$f(u) = \begin{cases} 1, & \alpha v u > \theta \\ 0, & \alpha v u \leq \theta \end{cases}$$

$$y = f(u)$$



Διαδικασία Διαμόρφωσης Βαρών σε ένα Νευρώνα Perceptron:

- Στόχος η διαμόρφωση των βαρών (W_i) των εισόδων (x_i) ώστε η παραγόμενη έξοδος (y_i) να είναι ίση με την επιθυμητή έξοδο (d_i) για κάθε διάνυσμα t (παράδειγμα εκπαίδευσης) εισαγωγής
- Ελέγχουμε το σφάλμα (e) που έχει η παραγόμενη έξοδος (y_i) για κάθε διάνυσμα (t):

$$e(t) = |d(t) - y(t)|, \quad e(t) \in \{0,1\}$$

- Εάν το σφάλμα είναι ίσο με το μηδέν ($e=0$) τότε δεν κάνουμε καμία τροποποίηση των βαρών του νευρώνα
- Εάν το σφάλμα είναι ίσο με ένα ($e=1$) τότε για κάθε διάνυσμα έχουμε:

$$\Delta w_i(t) = x_i(t) \cdot e(t) \cdot n$$

$$\Delta b(t) = e(t) \cdot n$$

Όπου n ο ρυθμός μάθησης του νευρώνα και παίρνει τιμές στο διάστημα $[0,1]$

Παράδειγμα εφαρμογής ενός Νευρώνα Perceptron:

Έστω πως έχουμε τα δεδομένα ενός προβλήματος, που παρουσιάζονται στον παρακάτω πίνακα:

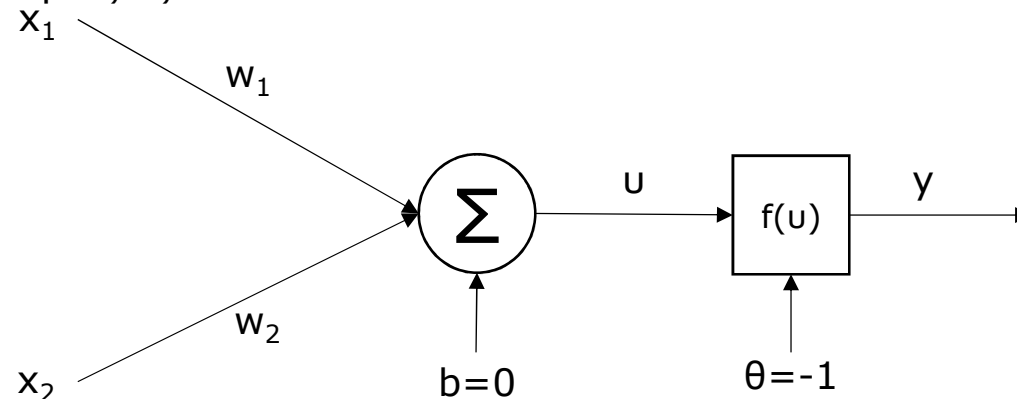
		T
x	x ₂	d
1	0	1
0	1	0

- i. Σχεδιάστε τον αντίστοιχο νευρώνα με αρχικές τιμές βαρών $w_1=1$ και $w_2=2$, τιμή κατωφλίου $b=0$, τιμή κατωφλιού $\theta=-1$ και ρυθμό μάθησης $\eta=0,5$
- ii. Περιγράψτε τη λειτουργία του αντιστοίχου νευρώνα
- iii. Ελέγξτε αν σε κάθε περίπτωση αν είναι απαραίτητη η τροποποίηση των βαρών του νευρώνα και αν είναι απαραίτητη περιγράψτε τη διαδικασία τροποποίησης καθώς και τις τελικές τιμές των βαρών για κάθε περίπτωση

Παράδειγμα εφαρμογής ενός Νευρώνα Perceptron (Σχεδιασμός Νευρώνα):

T		
x	x_2	d
1	0	1
0	1	0

Όπως παρατηρούμε στα δεδομένα μας οι επιθυμητές τιμές εξόδου ($d_1=1$ και $d_2=0$) είναι δυαδικής μορφής. Επομένως το συγκεκριμένο πρόβλημα ταξινόμησης μπορεί να λυθεί με ένα νευρώνα Perceptron δύο εισόδων και μιας εξόδου:



Παράδειγμα εφαρμογής ενός Νευρώνα Perceptron (Λειτουργία του Νευρώνα):

$$v = \sum_{i=1}^2 w_i \cdot x_i + b$$

$$y = f(v)$$

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

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ και } D = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$b=0$$

$$\theta=-1$$

$$n=0,5$$

1^η Περίπτωση (t_1) από τα δεδομένα μας:

$$x_1=1$$

$$w_1=1$$

$$b=0$$

$$x_2=0$$

$$w_2=2$$

$$\theta=-1$$

$$d_1=1$$

$$v_1 = \sum_{i=1}^2 w_i \cdot x_i + b = 1 \cdot 1 + 0 \cdot 2 + 0 \Rightarrow v_1 = 1$$

$$f(v_1) = f(1) = 1. \text{ αφού } 1 > -1$$

2^η Περίπτωση (t_1) από τα δεδομένα μας:

$$x_1=0$$

$$w_1=1$$

$$b=0$$

$$x_2=1$$

$$w_2=2$$

$$\theta=-$$

$$1$$

$$d_1=0$$

$$v_2 = \sum_{i=1}^2 w_i \cdot x_i + b = 1 \cdot 0 + 1 \cdot 2 + 0 \Rightarrow v_2 = 2$$

$$f(v_2) = f(2) = 1. \text{ αφού } 2 > -1$$



Παράδειγμα εφαρμογής ενός Νευρώνα Perceptron (Έλεγχος Σφαιλμάτων):

1^η Περίπτωση (t_1) από τα δεδομένα μας:

$$\begin{array}{lll} x_1=1 & w_1=1 & b=0 \\ x_2=0 & w_2=2 & \theta=-1 \\ d_1=1 & u_1=1 & y_1=1 \end{array}$$

$$e(t) = |d(t) - y(t)| \Rightarrow e(t_1) = |d_1 - y_1| \Rightarrow e(t_1) = |1-1| \Rightarrow e(t_1) = 0$$

2^η Περίπτωση (t_2) από τα δεδομένα μας:

$$\begin{array}{lll} x_1=0 & w_1=1 & b=0 \\ x_2=1 & w_2=2 & \theta=-1 \\ d_2=0 & u_2=2 & y_2=1 \end{array}$$

$$e(t) = |d(t) - y(t)| \Rightarrow e(t) = |d_2 - y_2| \Rightarrow e(t) = |0-1| \Rightarrow \mathbf{e(t_2) = 1}$$

Παράδειγμα εφαρμογής ενός Νευρώνα Perceptron (Τροποποίηση Βαρών):

Συμπεράσματα από τον έλεγχο των σφαλμάτων:

- Όπως παρατηρούμε κατά την εισαγωγή των δεδομένων της πρώτης περίπτωσης (t_1) του προβλήματος με τις αρχικές τιμές των βαρών του νευρώνα είχαμε μηδενικό σφάλμα ($e(t_1)=0$), επομένως σε αυτή τη περίπτωση δεν απαιτείται κάποια τροποποίηση των αρχικών τιμών των βαρών του νευρώνα
- Όμως, κατά της εισαγωγή των δεδομένων της δεύτερης περίπτωσης (t_2) του προβλήματος με τις αρχικές τιμές των βαρών του νευρώνα παρατηρούμε πως η παραγόμενη έξοδος ($y_2=1$) διαφέρει από την επιθυμητή έξοδο ($d_2=0$) κατά μια μονάδα ($e(t_2)=1$). Επομένως σε αυτή τη περίπτωση θα τροποποιήσουμε τα βάρη του νευρώνα καθώς και τη τιμή του κατωφλίου ($b=0$) σύμφωνα με το ρυθμό μάθησης που μας έχει δοθεί ($\eta=0,5$) και σύμφωνα με τις εξισώσεις:

$$\Delta w_i(t) = x_i(t) \cdot e(t) \cdot \eta$$

$$\Delta b(t) = e(t) \cdot \eta$$



Παράδειγμα εφαρμογής ενός Νευρώνα Perceptron (Τροποποίηση Βαρών):

Συναρτήσεις Τροποποίησης Βαρών και Κατώφλιου b:

$$\Delta w_i(t) = x_i(t) \cdot e(t) \cdot n$$

$$\Delta b(t) = e(t) \cdot n$$

$$\Delta w_1(t_2) = x_1(t_2) \cdot e(t_2) \cdot n \Rightarrow \Delta w_1(t_2) = 0 \cdot 1 \cdot 0,5 \Rightarrow \Delta w_1(t_2) = 0$$

επομένως το νέο βάρος για την είσοδο X_1 θα είναι $w_{1(new)} = w_{1(old)} - \Delta w_1(t_2) = 1 - 0 \Rightarrow \mathbf{w_{1(new)} = 1}$

$$\Delta w_2(t_2) = x_2(t_2) \cdot e(t_2) \cdot n \Rightarrow \Delta w_2(t_2) = 1 \cdot 1 \cdot 0,5 \Rightarrow \Delta w_2(t_2) = 0,5$$

επομένως το νέο βάρος για την είσοδο X_2 θα είναι $w_{2(new)} = w_{2(old)} - \Delta w_2(t_2) = 2 - 0,5 \Rightarrow \mathbf{w_{2(new)} = 1,5}$

$$\Delta b(t_2) = e(t) \cdot n = 1 \cdot 0,5 \Rightarrow \Delta b(t_2) = 0,5$$

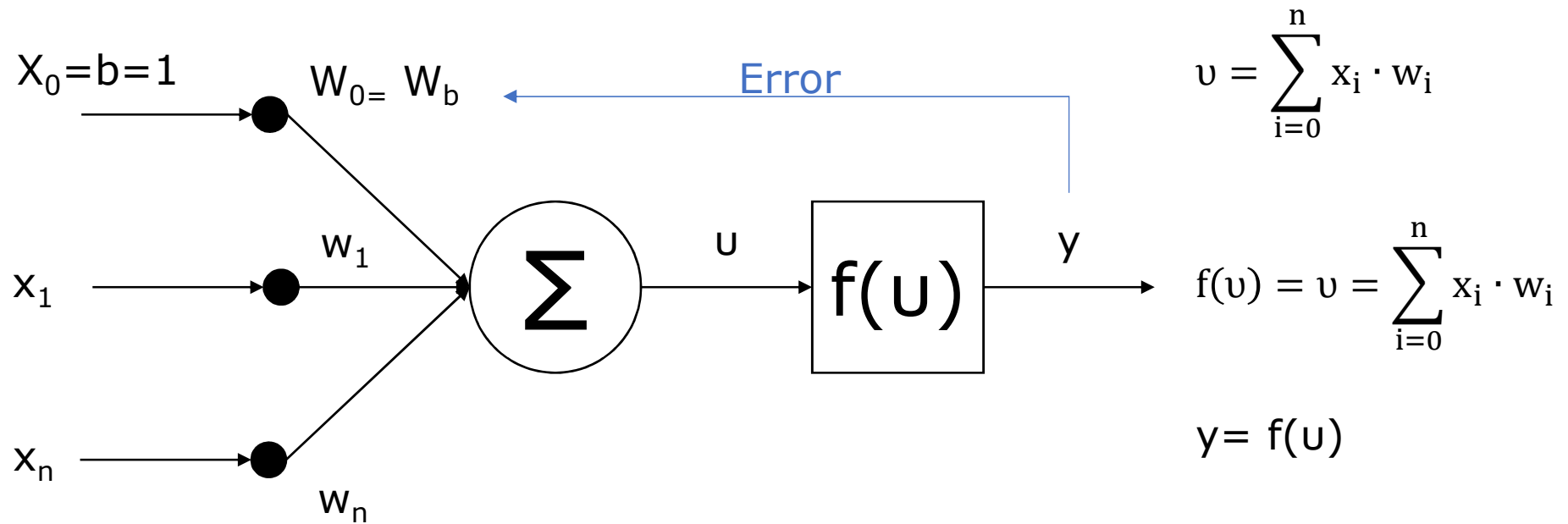
Επομένως το νέο κατώφλι (b) του νευρώνα θα είναι $b_{new} = b_{old} - \Delta b(t_2) = 0 - 0,5 \Rightarrow \mathbf{b_{new} = -0,5}$



Νευρώνας ADALINE: ADaptive LINear Element

- Προτάθηκε από τους Widrow και Hoff το 1962
- Είναι μια ελαφρώς τροποποιημένη έκδοση του νευρώνα Perceptron
- Δέχεται πραγματικές τιμές ως εισόδους
- Έχει συνοπτικά βάρη
- Έχει ως συνάρτηση ενεργοποίησης τη γραμμική συνάρτηση
- Μπορεί να χρησιμοποιηθεί για προβλήματα ταξινόμησης





Διαδικασία Διαμόρφωσης Βαρών σε ένα Νευρώνα ADALINE:

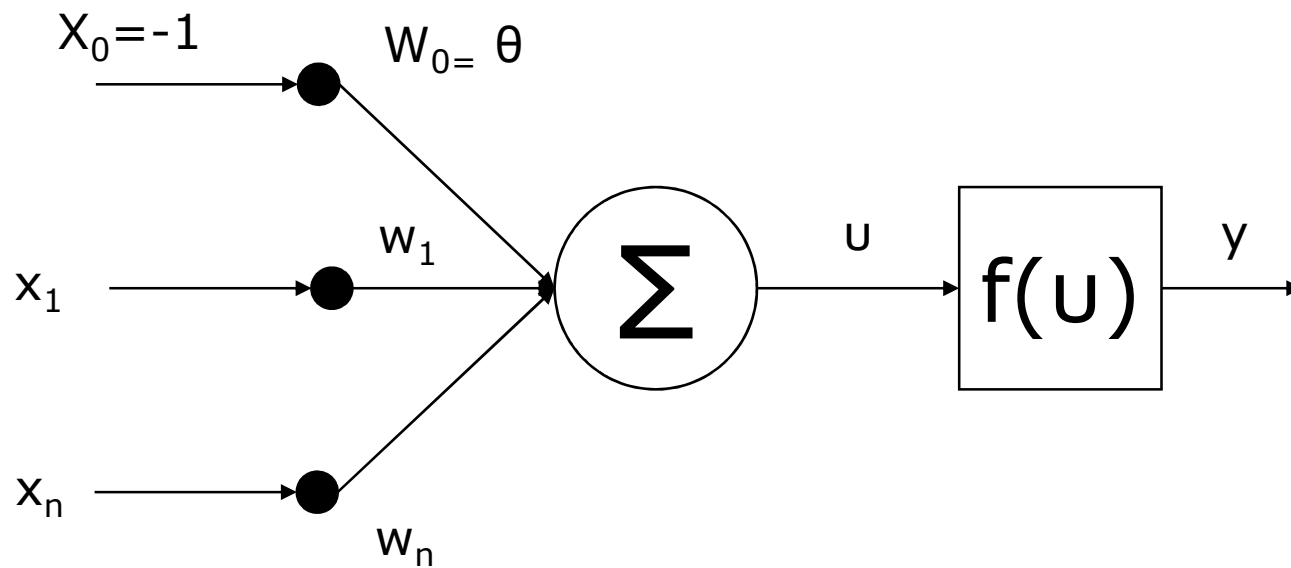
- Στόχος η διαμόρφωση των βαρών (w_i) των εισόδων (x_i) ώστε η παραγόμενη έξοδος (y_i) να είναι ίση με την επιθυμητή έξοδο (d_i) για κάθε διάνυσμα t (παράδειγμα εκπαίδευσης) εισαγωγής
- Ελέγχουμε το σφάλμα (e) που έχει η παραγόμενη έξοδος (y_i) για κάθε διάνυσμα (t):

$$e(t) = d(t) - \sum_{i=0}^n x_i \cdot w_i, e(t) \in \mathbb{R}$$

- Εάν το σφάλμα είναι διάφορο του μηδενός τότε βάση του κανόνα δ έχουμε:

$$w_i(t+1) = w_i(t) + n \cdot x_i(t) \cdot e(t)$$

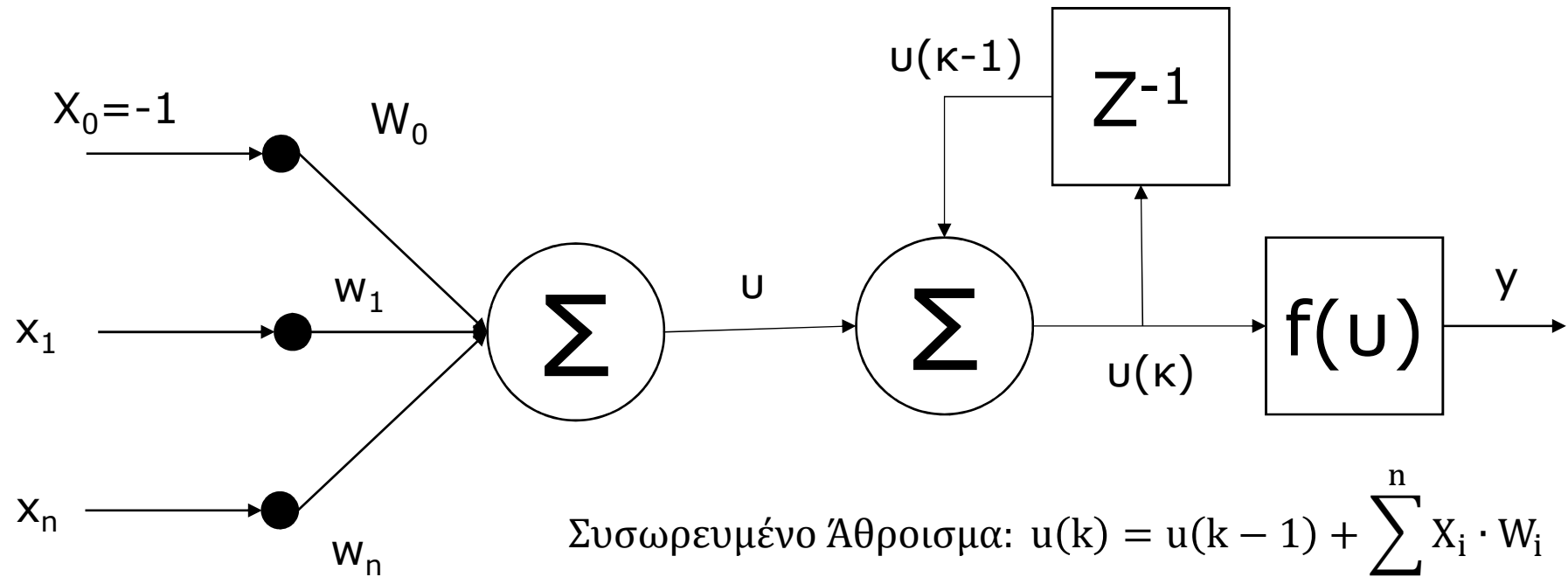
Όπου n ο ρυθμός μάθησης του νευρώνα και παίρνει τιμές στο διάστημα $[0,1]$



$$X = [X_1, X_2, \dots, X_n]^T$$

$$W = [W_1, W_2, \dots, W_n]^T$$

$$y = f\left(\sum_{i=0}^n X_i \cdot W_i\right)$$



Συσσωρευμένο Άθροισμα: $u(k) = u(k - 1) + \sum_{i=0}^n X_i \cdot W_i$

Δείκτης Διακριτού Χρόνου: k



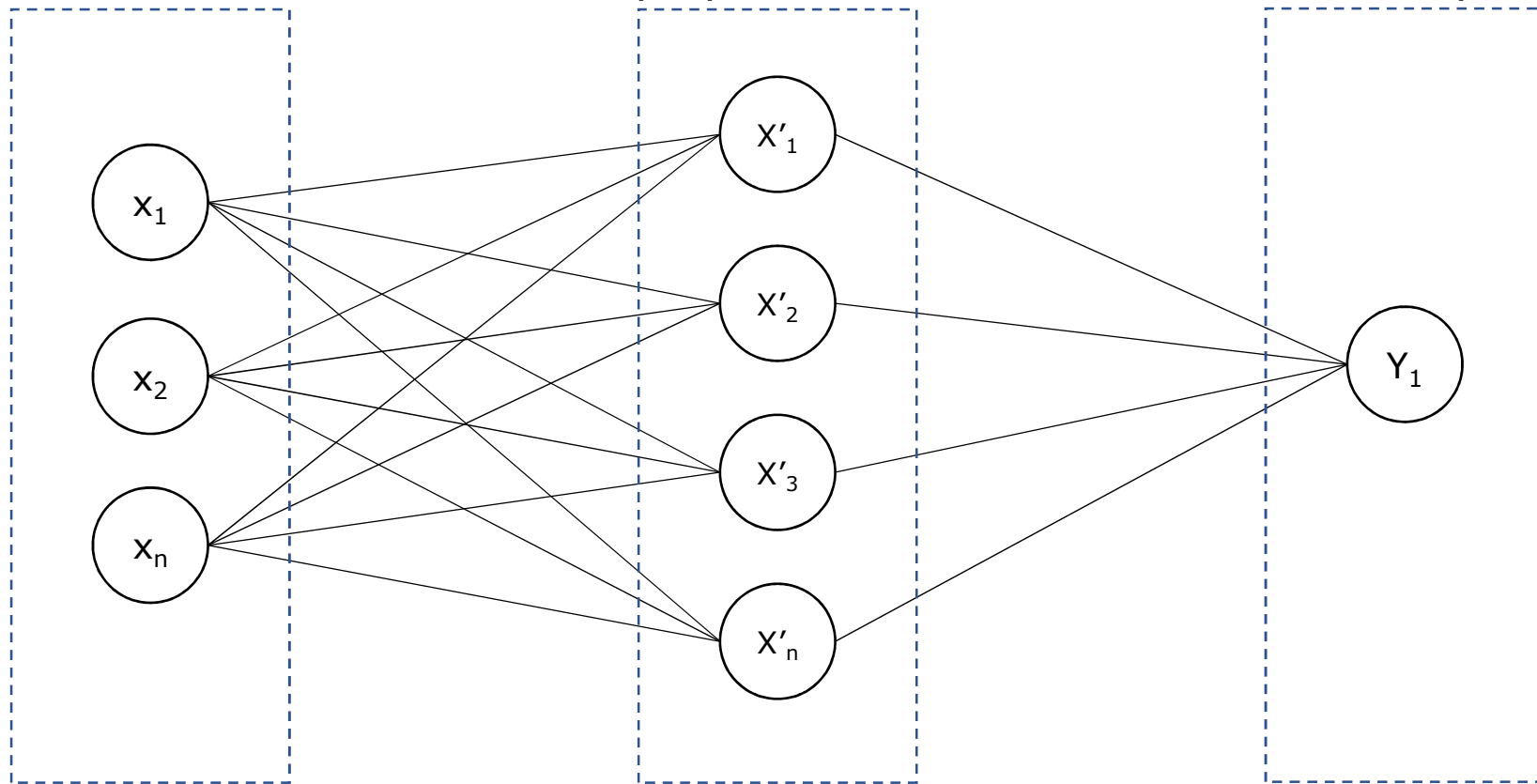
Αρχιτεκτονικές ΤΝΔ

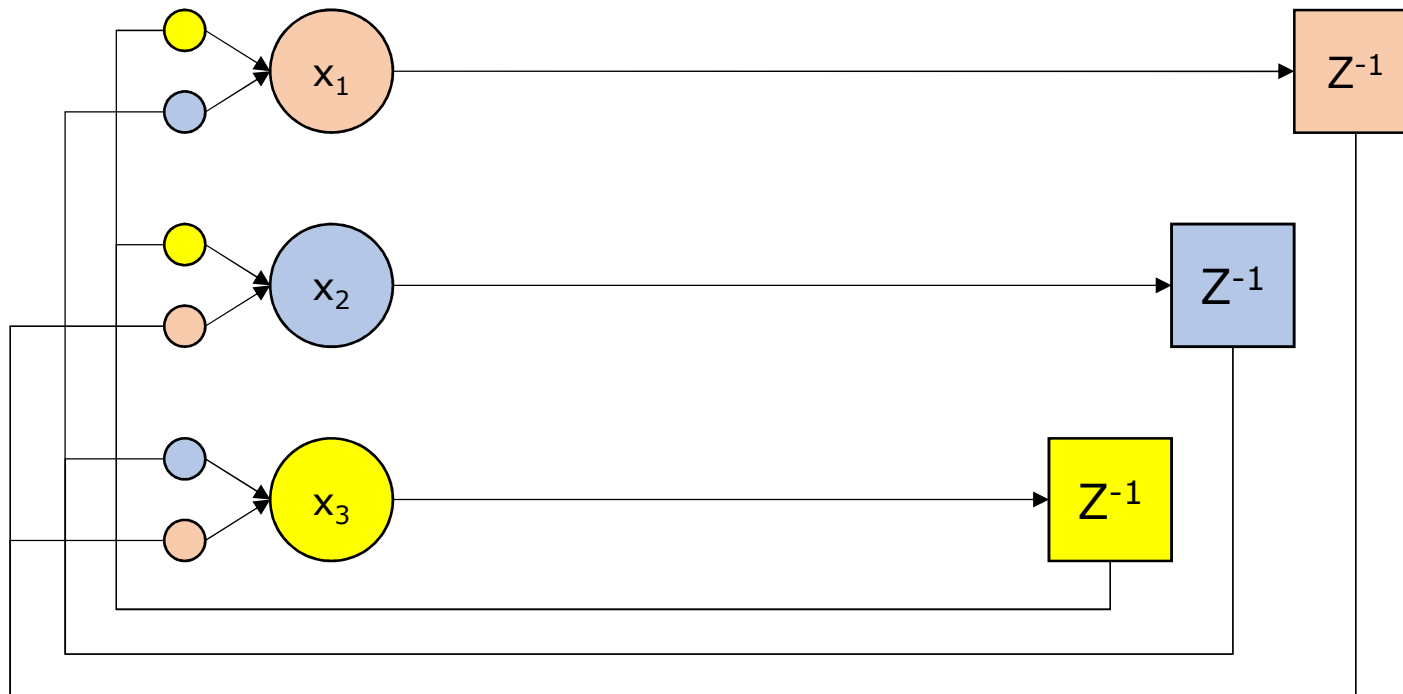


Επίπεδο Εισόδων

Κρυφό Επίπεδο

Επίπεδο Εξόδων





Μοντέλα Τεχνητών Νευρωνικών Δικτύων

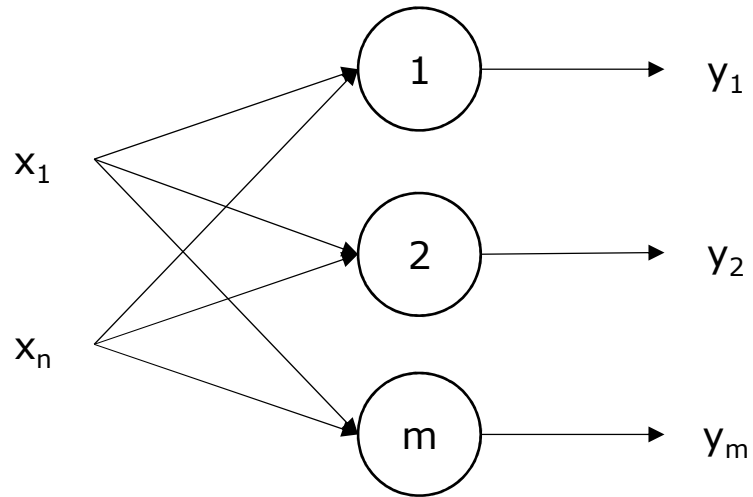


SLP

MLP

RBF

Μονοστρωματικό ΤΝΔ (SLP)



$$y_i = f\left(\sum_{i=1}^n w_{ij} \cdot x_i + \theta_i\right)$$

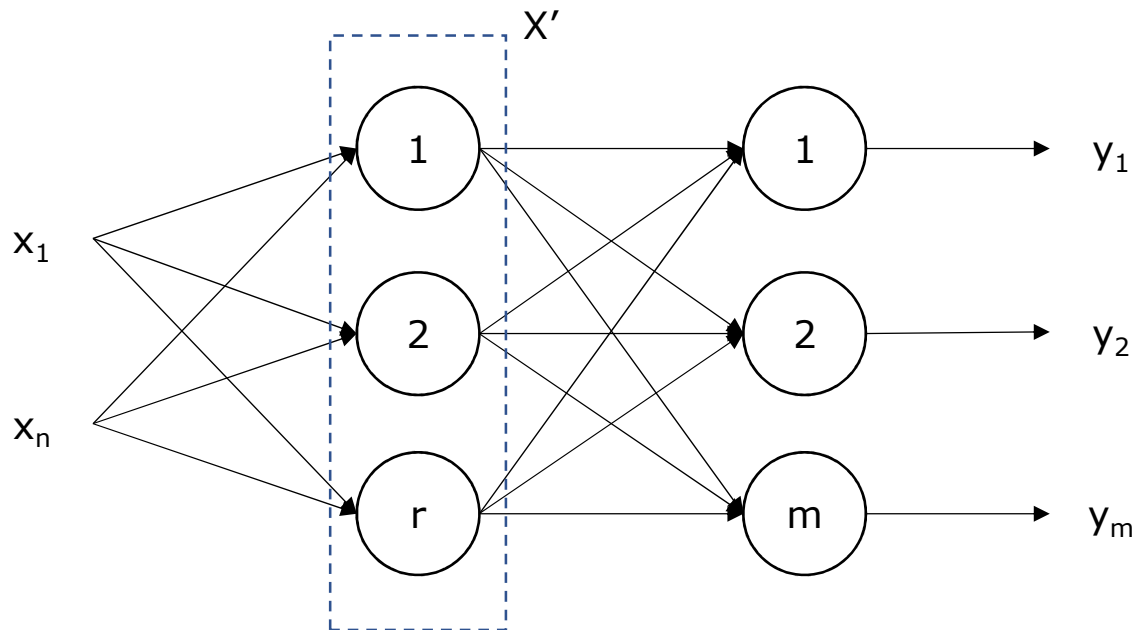


SLP

MLP

RBF

Πολυστρωματικό ΤΝΔ (MLP)

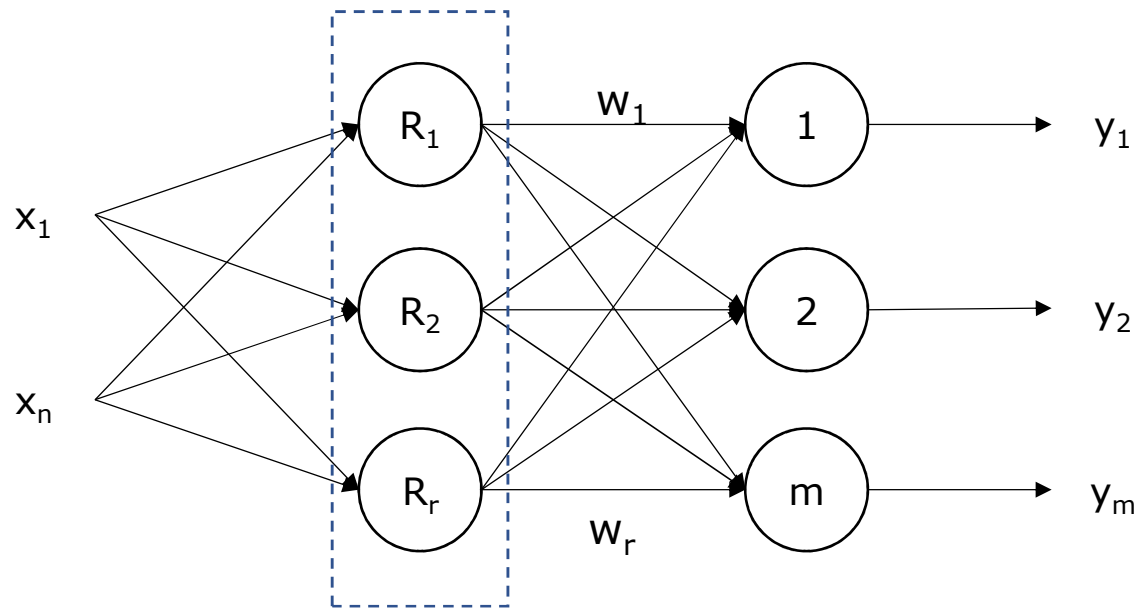


$$x'_i = f\left(\sum_{i=1}^n w_{ij} \cdot x_i + \theta_i\right)$$

$$y_i = f\left(\sum_{i=1}^r w'_{ij} \cdot x'_i + \theta_i\right)$$



ΤΝΔ με Ακτινικές Συναρτήσεις Βάσης(RBF)



$$y_i(\vec{x}) = \sum_{i=1}^m w_i \cdot R_i(\vec{x})$$

με κανονικοποιημένη έξοδο:

$$y_i(\vec{x}) = \frac{\sum_{i=1}^m w_i \cdot R_i(\vec{x})}{\sum_{i=1}^m R_i(\vec{x})}$$

Τοπική συνάρτηση βάσης ακτινικού τύπου (Gauss):

$$R_i(\vec{x}) = e^{-\frac{\|\vec{x} - c_i\|^2}{\sigma_i^2}}$$

Η συνάρτηση ακτινικού τύπου έχει κέντρο c και εύρος σ



ΤΝΔ με Ακτινικές Συναρτήσεις Βάσης(RBF)

- Είναι εντελώς όμοια με ένα δίκτυο MLP
- Είναι δίκτυα δυο στρωμάτων (χωρίς να υπολογίζουμε το στρώμα εισόδου), όπου το κρυφό στρώμα περιέχει την ακτινική συνάρτηση ενεργοποίησης και το στρώμα εξόδου
- ένα γραμμικό νευρώνα
- Η διαφορά των RBF και MLP είναι ότι ένα RBF πραγματοποιεί μια τοπική απεικόνιση ενώ ένα MLP πραγματοποιεί μια συνολική απεικόνιση
- Μπορούν να προσεγγίσουν οποιαδήποτε συνεχή συνάρτηση
- Είναι ισοδύναμα με μηχανές ασαφούς λογικής
- Αποτελούν μια εναλλακτική πρόταση μοντελοποίησης των ΤΝΔ



SLP

MLP

RBF

Εκπαίδευση RBF

- Κάθε ακτινικός νευρώνας αντιστοιχεί σε ένα παράδειγμα εκπαίδευσης.
- Δεν υπάρχουν συνοπτικά βάρη μεταξύ εισόδων και κρυφού στρώματος
- Στην πρώτη φάση εκπαίδευσης καθορίζονται τα κέντρα των τοπικών συναρτήσεων βάσης
- Στη δεύτερη φάση εκπαίδευσης καθορίζουμε τις συνδέσεις (συναπτικά βάρη) στο δεύτερο επίπεδο χρησιμοποιώντας τους αλγορίθμους βελτιστοποίησης βαθμίδας ή των ελαχίστων τετραγώνων



Σχεδιασμός RBF

- **Επιλογή των Κέντρων c των Κρυφών Νευρώνων**
 - ✓ Κάθε πρότυπο και ένα κέντρο
 - ✓ Τυχαία επιλογή προτύπων για να γίνουν κέντρα
 - ✓ Μέθοδοι ομαδοποίησης (clustering: K-μέσων, κ.α.)
- **Υπολογισμός της τιμής του εύρους σ των Γκαουσιανών**
 - ✓ Δίνουμε σε όλους τους νευρώνες το ίδιο εύρος $\sigma = \frac{D}{\sqrt{2P}}$, όπου D είναι η απόσταση μεταξύ των πιο απομακρυσμένων κέντρων και P το σύνολο των προτύπων
- **Υπολογισμός των Συναπτικών Βαρών**
 - ✓ Βασίζεται στην επίλυση συστήματος P γραμμικών εξισώσεων με P αγνώστους
 - ✓ Εκπαίδευση με επίβλεψη (Generalized RBF)
- Η **Ανάκληση** του δικτύου RBF είναι όμοια με αυτή ενός δικτύου MLP



Η Λειτουργία ενός ΤΝΔ



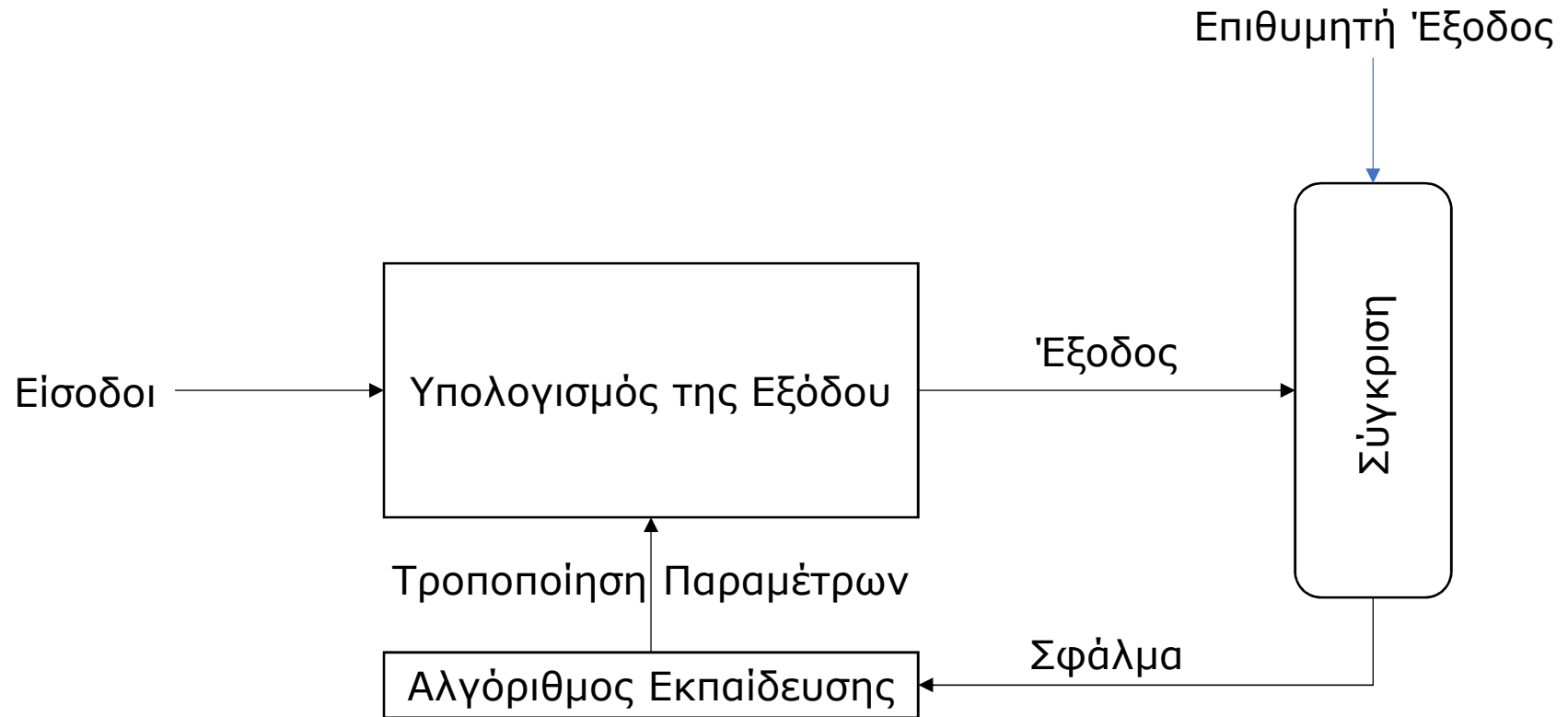
Εκπαίδευση του δικτύου:

- Παρουσιάζονται τα δεδομένα στο TND
- Το TND υπολογίζει την έξοδο
- Πραγματοποιείται η σύγκριση της παραγόμενης εξόδου με την επιθυμητή έξοδο
- Προκύπτει ένα σφάλμα μεταξύ της παραγόμενης και της επιθυμητής εξόδου
- Με τη χρήση του αλγόριθμου εκπαίδευσης του TND πραγματοποιείται η τροποποίηση των συνοπτικών βαρών με απώτερο σκοπό τη μείωση του σφάλματος

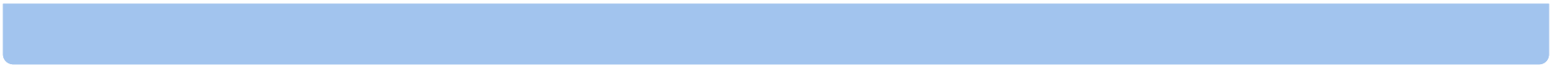
Χρήση του δικτύου:

- Παρουσίαση νέων δεδομένων στο TND
- Το εκπαιδευμένο (πια) TND υπολογίζει την έξοδο





Μάθηση Τεχνητών Νευρωνικών Δικτύων



Μάθηση (Learning) Τεχνητών Νευρωνικών Δικτύων:

- Διαδικασία τροποποίησης των Βαρών των συνδέσεων ενός Τεχνητού Νευρωνικού Δικτύου
- Πραγματοποιείται ανά Κύκλους Εκπαίδευσης-Αξιολόγησης, που ονομάζονται Εποχές (Epoches) και χαρακτηρίζεται τόσο από τη χρήση των δεδομένων όσο και από την ύπαρξη της επίβλεψης από τον άνθρωπο



Σύμφωνα με τη χρήση των δεδομένων:

- Μάθηση Δέσμης (Batch Learning)

Η τροποποίηση των βαρών πραγματοποιείται μετά από την ολική παρουσίαση (στο ΤΝΔ) των δεδομένων Εκπαίδευσης

- Επαυξητική Μάθηση (Incremental Learning)

Η τροποποίηση των Βαρών πραγματοποιείται ανά διάνυσμα εκπαίδευσης

- Συνδυασμός Επαυξητικής και Δέσμης

☀ Η επικρατέστερη μέθοδος είναι η Μάθηση Δέσμης

Σύμφωνα με την επίβλεψη της διαδικασίας της μάθησης:

- Επιβλεπόμενη Μάθηση (Supervised Learning)
- Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)
- Βαθμολογημένη Μάθηση (Graded Learning)

💡 Στις περισσότερες εφαρμογές ΤΝΔ πραγματοποιείται Επιβλεπόμενη Μάθηση αυτών



Επιβλεπόμενη Μάθηση (Supervised Learning) ΤΝΔ:

- Τα δεδομένα των προτύπων εκπαίδευσης του ΤΝΔ είναι χωρισμένο σε ζεύγη της μορφής Είσοδοι-Επιθυμητή Έξοδος
- Το σύστημα εποπτεύεται από τον άνθρωπο και συσχετίζει τις εισόδους με την έξοδο
- Ελαχιστοποίηση του σφάλματος (error) μεταξύ παραγόμενης και επιθυμητής εξόδου
- Χρησιμοποιείται, συνήθως, σε προβλήματα Ταξινόμησης (Classification), Πρόγνωσης (Prediction) και Διερμηνείας (Interpretation)



Μη Επιβλεπόμενη Μάθηση (Supervised Learning) ΤΝΔ:

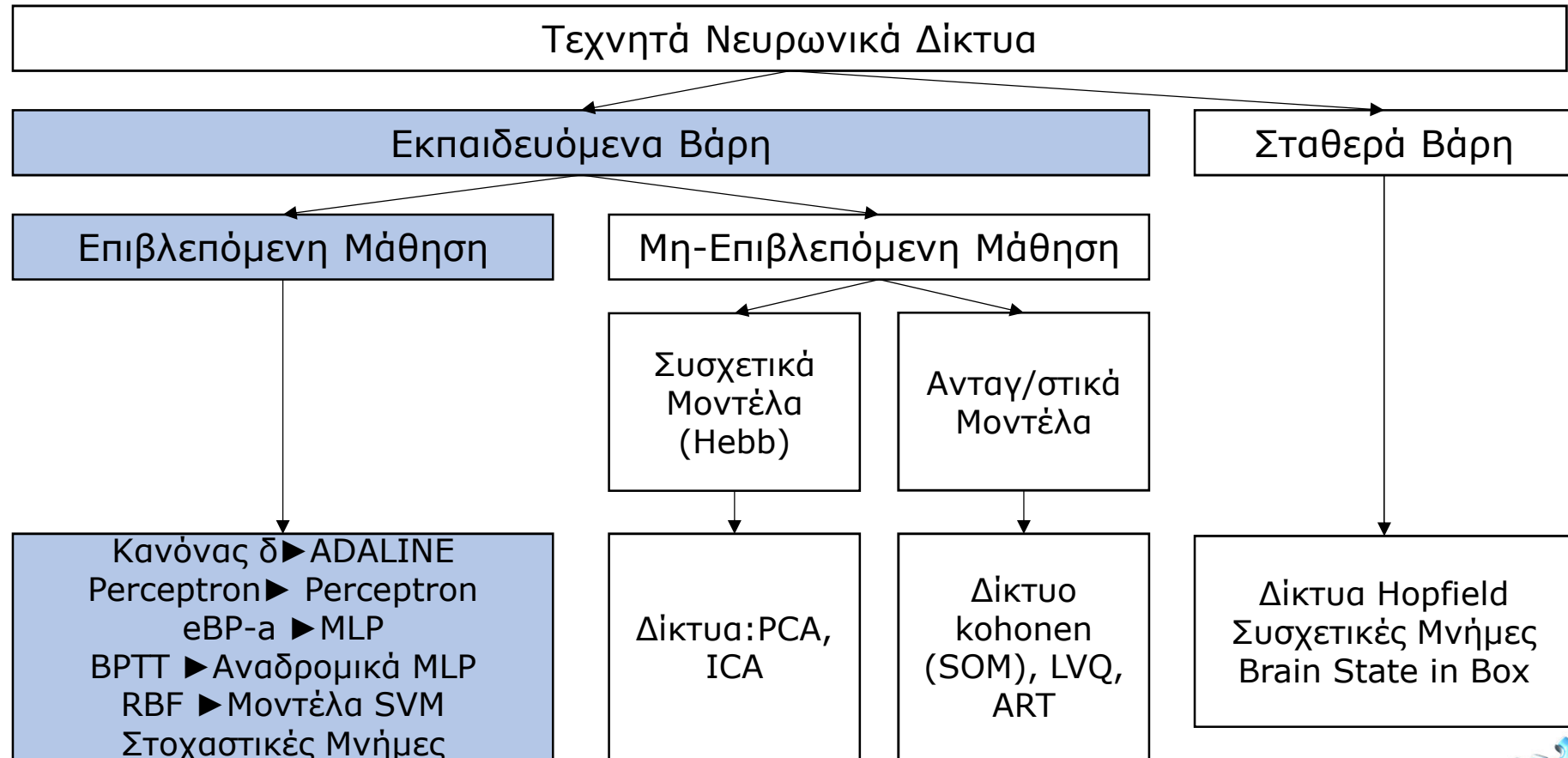
- Τα πρότυπα εκπαίδευσης περιλαμβάνουν μόνο τα Διανύσματα Εισόδου
- Εξαγωγή βασικών δομικών ιδιοτήτων των δεδομένων εισόδου (λ.χ. εύρεση ομάδων)
- Χρησιμοποιείται, συνήθως, σε προβλήματα Ανάλυσης Συσχετισμών (Association Analysis) και Ομαδοποίησης (Clustering)



Βαθμολογημένη Μάθηση (Graded Learning):

- Τα πρότυπα εκπαίδευσης περιλαμβάνουν μόνο τα Διανύσματα Εισόδου
- Το σύστημα εποπτεύεται από τον άνθρωπο και συσχετίζει τις εισόδους με την έξοδο
- Παρατήρηση και αξιολόγηση της συμπεριφοράς του ΤΝΔ βάση του δείκτη συμπεριφοράς
- Δείκτης συμπεριφοράς → Σήμα ενίσχυσης (reinforcement signal) → Επιβράβευση (Reward) ή Επιβολή Ποινής (Penalization)
- Χρησιμοποιείται, συνήθως, στη ρομποτική (robotics) και στα ηλεκτρονικά παιχνίδια (games)





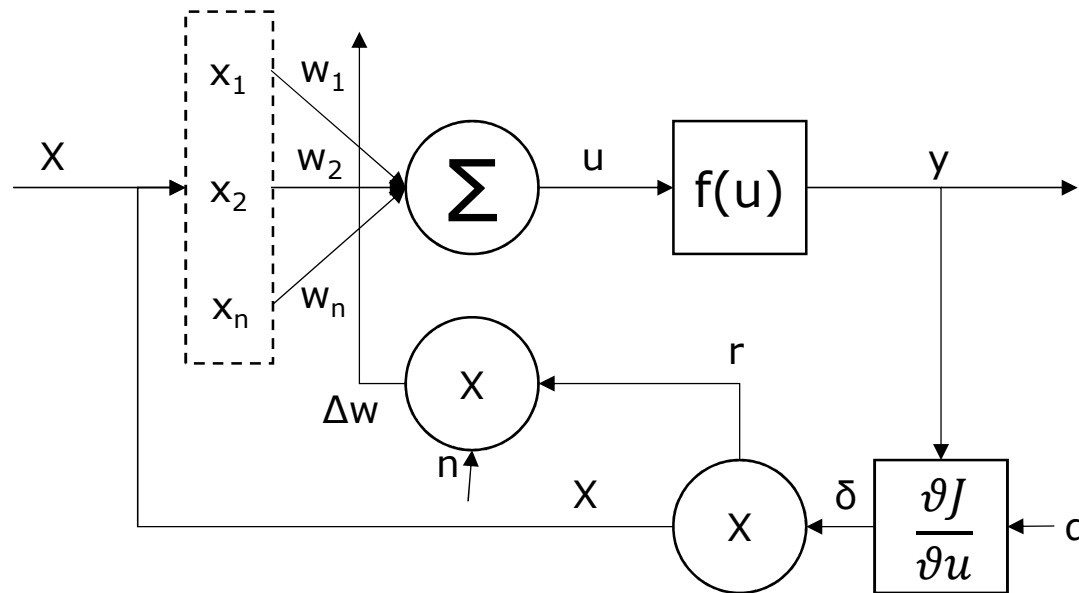
Κανόνας δ



Delta Rule:

- Windrow-Hoff Learning Rule ή Least Mean Square Learning Rule
- Παραλλαγή του Κανόνα Hebb
- Συνεχή τροποποίηση των βαρών εισόδου με απώτερο σκοπό τη μείωση του σφάλματος μεταξύ επιθυμητής και παραγόμενης εξόδου ενός νευρώνα
- Διόρθωση των βαρών σε κάθε περίπτωση εκπαίδευσης μέσω της διαδικασίας των ελαχίστων τετραγώνων
- Δεν μπορεί να εφαρμοστεί σε ΤΝΔ με κρυφά επίπεδα





$$J = \frac{1}{2} \cdot (d - y)^2$$

$$\delta = -\frac{\partial J}{\partial u} = -\frac{\partial J}{\partial y} \cdot \frac{\partial y}{\partial u} \Rightarrow$$

$$\Rightarrow \delta = (d - y) \cdot f'(u)$$

$$\Rightarrow \delta = e \cdot f'(u)$$

$\delta = e \cdot f'(u)$: Για παραγωγίσιμη συνάρτηση ενεργοποίησης

$\delta = e$: Για μη παραγωγίσιμη συνάρτηση ενεργοποίησης

$$\left. \begin{array}{l} r = \delta \cdot X \\ \Delta \chi = n \cdot r \end{array} \right\} \Delta w = n \cdot \delta \cdot X$$

u: Δικτυακή Διέγερση του Νευρώνα

δ: Σφάλμα ή Τοπική Βαθμίδα ή Παράγωγος του Κόστους ως προς τη Δικτυακή Διέγερση

r: Σήμα Μάθησης

n: Ρυθμός Μάθησης

Αλγόριθμος Οπισθοδιάδοσης του Σφάλματος



Αλγόριθμος Ανάστροφης Μετάδοσης του Σφάλματος:

- Βασίζεται στο κανόνα δ για συνεχείς συναρτήσεις ενεργοποίησης
- Δημοφιλής για τα πολυ-επίπεδα τεχνητά νευρωνικά δίκτυα
- Κεντρική Ιδέα: Καθορισμός και ελαχιστοποίηση του σφάλματος σε κάθε νευρώνα για κάθε στρώμα του ΤΝΔ
- Η αναπροσαρμογή των βαρών γίνεται από το στρώμα εξόδου προς το στρώμα είσοδού
- Ακολουθείται η διαδικασία βελτιστοποίησης επικλινούς καθόδου
- Η απόδοση του δικτύου πραγματοποιείται βάση προτύπων αξιολόγησης (Validation Set)

Προβλήματα:

Παρουσιάζονται τοπικά ελάχιστα

Ένα ή περισσότερα βάρη έχουν σταθερά υψηλές απόλυτες τιμές και δεν τροποποιούνται σημαντικά σε κάθε διόρθωση (Network paralysis)

Δεν μπορεί να υπάρξει εγγύηση γενίκευσης, ακόμη και με ελάχιστο σφάλμα δ

Η Διόρθωση των Συνοπτικών Βαρών βασίζεται στον Κανόνα δ:

$$\left\{ \begin{array}{l} \text{Διόρθωση} \\ \text{Βαρών} \\ \Delta \mathbf{w} \end{array} \right\} = \left\{ \begin{array}{l} \text{Ρυθμός} \\ \text{Μάθησης} \\ \mathbf{n} \end{array} \right\} \cdot \left\{ \begin{array}{l} \text{Τοπικη} \\ \text{Βαθμίδα} \\ \delta \end{array} \right\} \cdot \left\{ \begin{array}{l} \text{Σήμα Εισοδου} \\ \text{του Νευρώνα} \\ \mathbf{x} \end{array} \right\}$$



Τύποι Υπολογισμού της Τοπικής Βαθμίδας:

- Εάν ο Νευρώνας k είναι Κόμβος Εξόδου:

$$\delta_k(L) = (d_k - y_k) \cdot f'(u_k) = e_k \cdot f'(u_k)$$

- Εάν ο Νευρώνας i είναι Κόμβος του Κρυφού Επιπέδου I :

$$\delta_i(I) = f'(u_i) \cdot \sum_l \delta_k(L) \cdot w_{ki}$$

Τύποι Ενημέρωσης των Βαρών:

- Για Συνοπτικά Βάρη μεταξύ Εξόδου και Κρυφού Επιπέδου:

$$\Delta w_{ki}(L) = n \cdot \delta_k(L) \cdot y_i$$

- Για Συνοπτικά Βάρη μεταξύ Κρυφού Επιπέδου και Εισόδου:

$$\Delta w_{ji}(L) = n \cdot \delta_i(L) \cdot x_j$$

Μαθηματική Περιγραφή

j : Αριθμός Εισόδου

i : Αριθμός Κρυφού Κόμβου

k : Αριθμός Κόμβου Εξόδου

I : Αριθμός Επιπέδου

Τα σφάλματα μοιάζουν να πηγάζουν από το τελευταίο στρώμα και να προωθούνται προς τα πίσω, δηλαδή μέχρι το πρώτο στρώμα. Τα σφάλματα που προωθούνται προς τα πίσω χρησιμοποιούν τα ίδια βάρη του δικτύου που χρησιμοποιούνται και στην ανάκληση.

Αξιολόγηση των ΤΝΔ

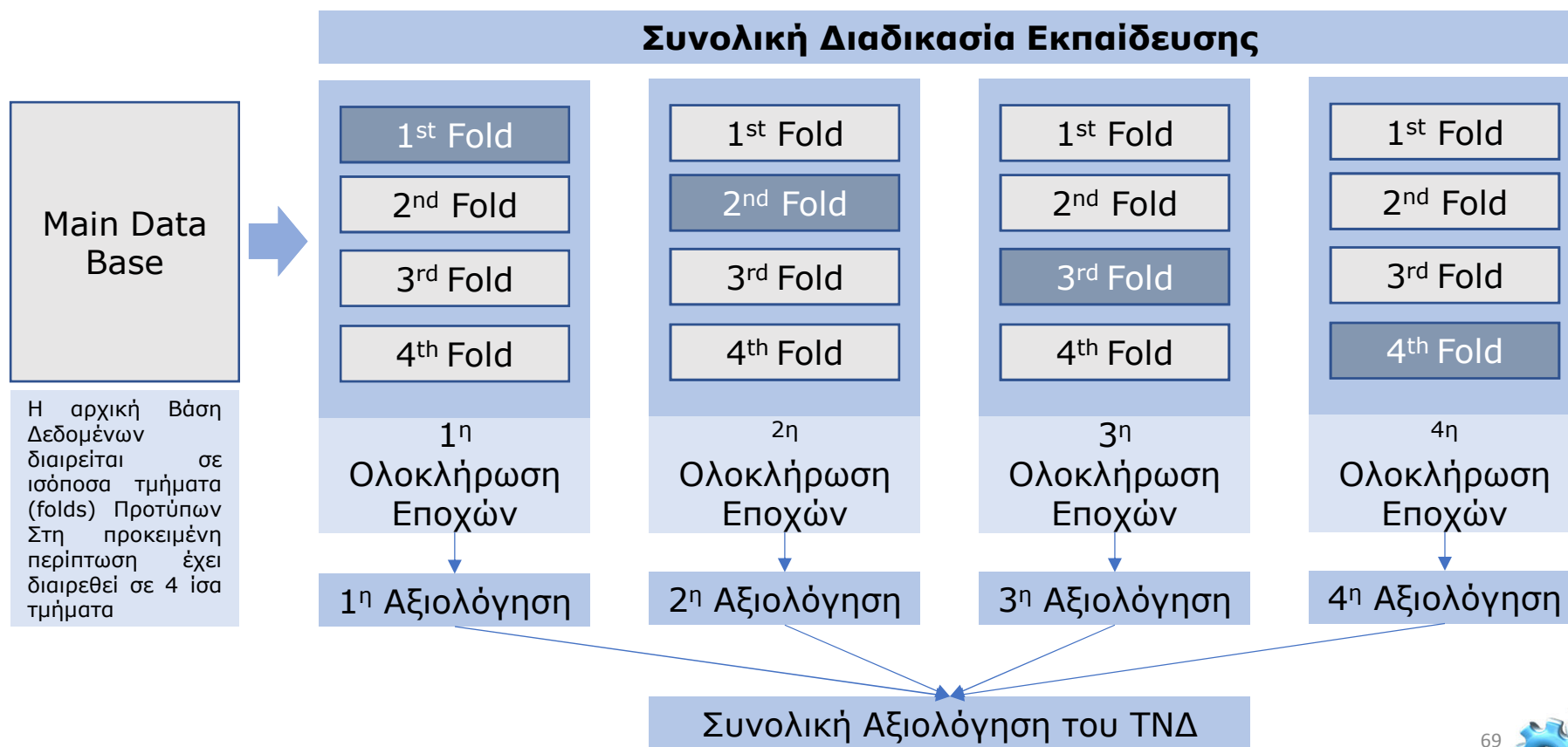


k-fold Cross Val.

Hold Out

Leave-1-Out

Διασταυρούμενη Επικύρωση k τμημάτων (k-fold Cross Validation):



Διασταυρούμενη Επικύρωση k τμημάτων (k-fold Cross Validation):

- Χωρίζουμε τη Βάση Δεδομένων σε k ($k=1,2,3,\dots,n$) ισόποσα τμήματα προτύπων
- Σε κάθε ολοκλήρωση των εποχών εκπαίδευσης το ΤΝΔ εκπαιδεύεται με τα $k-n$ τμήματα και αξιολογείται με το $k=1, 2,3,\dots,k$ τμήμα
- Συνήθως, η αξιολόγηση του ΤΝΔ προκύπτει από το μέσο σφάλμα (μεταξύ παραγόμενων και επιθυμητών εξόδων) σε κάθε ολοκλήρωση των εποχών εκπαίδευσης του.

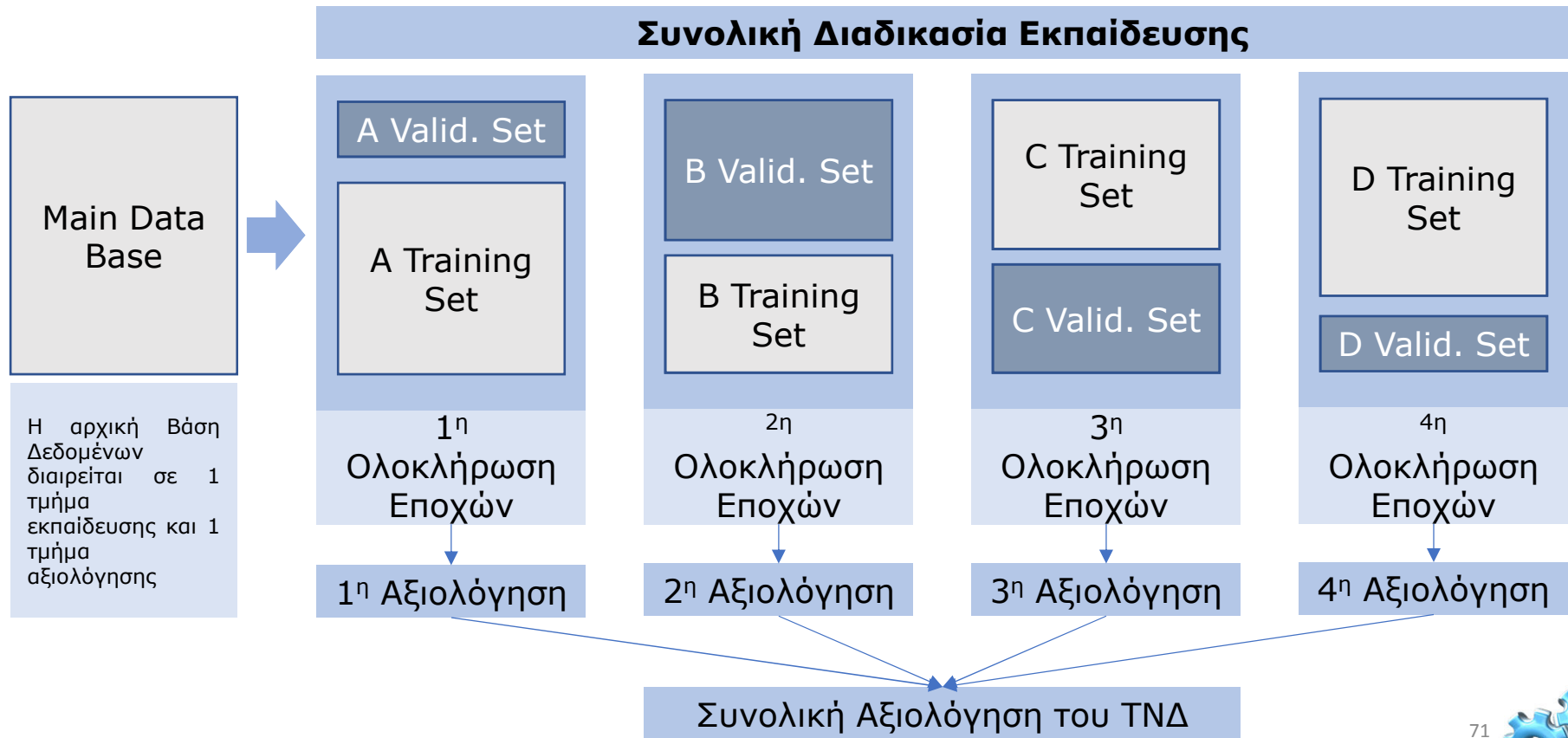


k-fold Cross Val.

Hold Out

Leave-1-Out

Πολλαπλή Hold Out:



k-fold Cross Val.

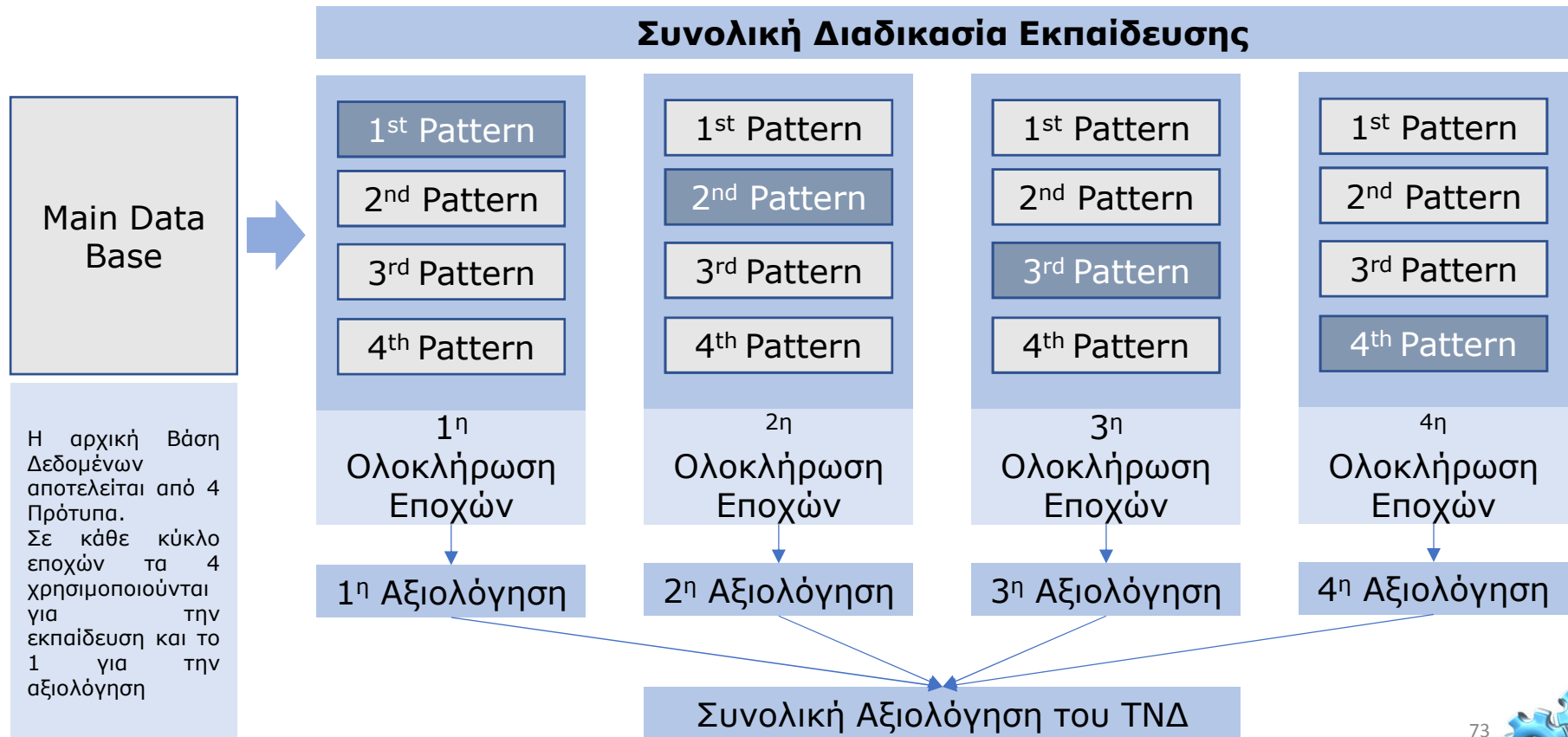
Hold Out

Leave-1-Out

Πολλαπλή Hold Out:

- Η αρχική Βάση Δεδομένων διαιρείται σε 1 τμήμα εκπαίδευσης και 1 τμήμα αξιολόγησης
- Υπάρχει πρόβλημα όταν έχουμε λίγα πρότυπα
- Μπορούμε να επαναλάβουμε τη διαδικασία Hold Out πολλές φορές και με τυχαίες διαιρέσεις της βάσης δεδομένων
- Η τελική εκτίμηση του σφάλματος γενίκευσης προκύπτει ως ο μέσος όρος των σφαλμάτων των επιμέρους εκπαιδεύσεων



Leave One Out (4 Πρότυπα):

Leave One Out (4 Πρότυπα):

- Ειδική περίπτωση της διαδικασίας k-fold cross validation για N πρότυπα και για $k=N$
- Η διαδικασία της εκπαίδευσης επαναλαμβάνεται N φορές
- Σε κάθε εκπαίδευση το ΤΝΔ εκπαιδεύεται με τα $N-1$ πρότυπα και αξιολογείται με το 1
- Η τελική εκτίμηση του σφάλματος γενίκευσης προκύπτει ως ο μέσος όρος των σφαλμάτων των επιμέρους εκπαιδεύσεων



Κριτήρια Αξιολόγησης Ταξινομητή



Ακρίβεια

Ευαισθησία

Προσδιοριστικότητα

Ακρίβεια (accuracy):

Λόγος των επιτυχών προβλέψεων επί του συνόλου των προτύπων

$$\frac{TP + TN}{TP + FP + FN + TN} = \frac{\text{Σωστές Προβλέψεις}}{\text{Σύνολο των Περιπτώσεων}}$$

Ακρίβεια

Ευαισθησία

Προσδιοριστικότητα

Ευαισθησία (sensitivity):

Λόγος των ορθών θετικών προβλέψεων, επί του συνόλου των θετικών προτύπων (True Positive Rate)

$$\frac{TP}{TP + FN} = \frac{\text{Ορθά Θετικά}}{\text{Θετικά}}$$

Προσδιοριστικότητα (specificity):

Λόγος των ορθών αρνητικών προβλέψεων, επί του συνόλου των αρνητικών προτύπων

$$\frac{TN}{TN + FP} = \frac{\text{Ορθο}\Phi \text{ Αρνητικ}\Phi}{\text{Αρνητικ}\Phi}$$

$$\text{False Positive Rate} = 1 - \text{Specificity} = 1 - \frac{TN}{TN + FP} = \frac{FP}{TN + FP} = \frac{\text{Εσφαλμ}\chi\text{να Θετικ}\Phi}{\text{Αρνητικ}\Phi}$$



Γενίκευση των ΤΝΔ



Με τον όρο Γενίκευση χαρακτηρίζουμε την αντίστοιχη της εκπαίδευσης ικανότητα, που έχει ένα ΤΝΔ στη σύγκλιση μεταξύ επιθυμητής και παραγόμενης εξόδου, σε άγνωστα (για αυτό) πρότυπα

- Χαρακτηρίζεται από το σφάλμα γενίκευσης
- Αποτελεί τη συμπερασματικό στάδιο της αξιολόγησης

Εκτίμηση της ικανότητας Γενίκευσης ενός ΤΝΔ:

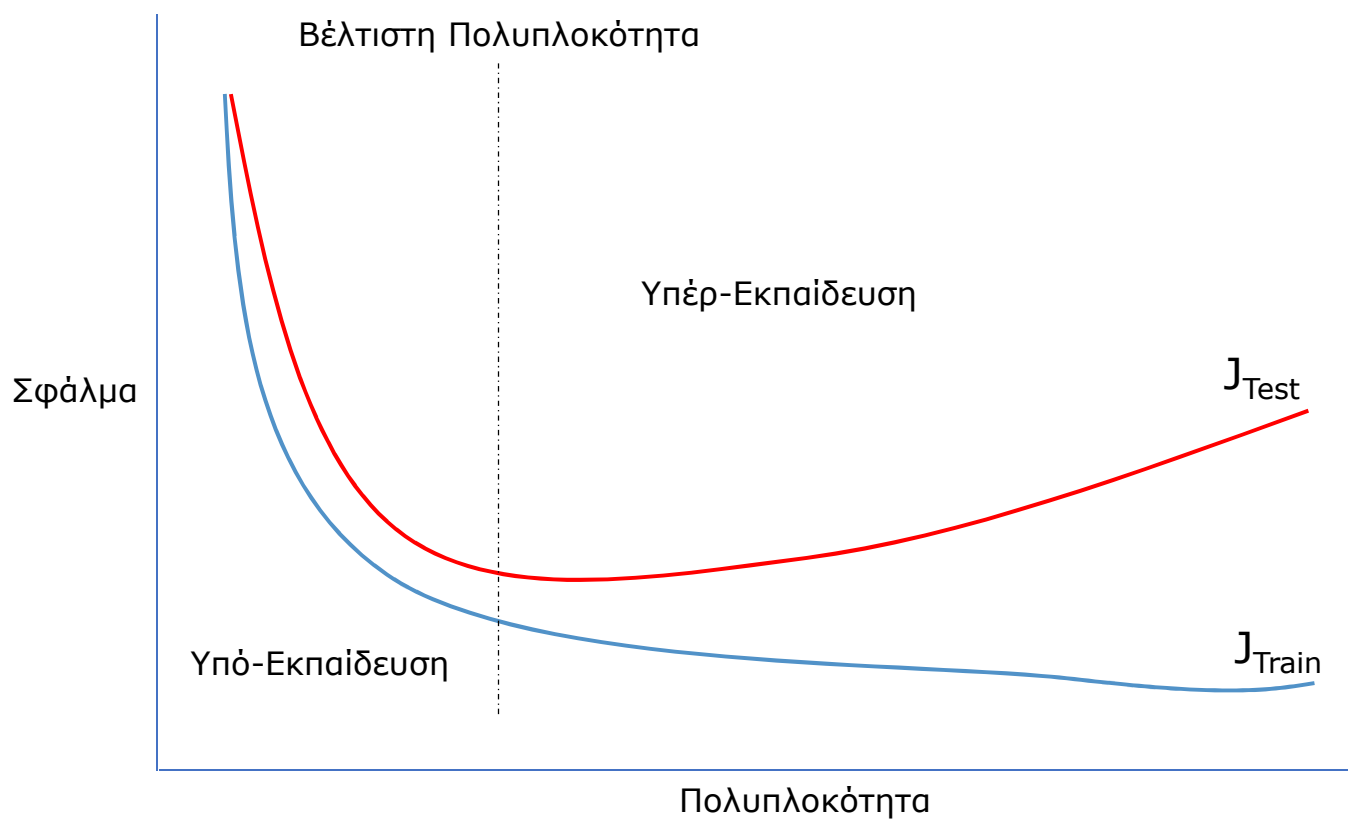
- Τα αρχικά δεδομένα χωρίζονται σε δύο τμήματα (ανάλογα με τη μέθοδο αξιολόγησης), τμήμα της εκπαίδευσης του ΤΝΔ (Train Set) και το τμήμα δοκιμής του ΤΝΔ (Test Set)
- Εκτιμάται το σφάλμα εκπαίδευσης J_{train} και το σφάλμα δοκιμής J_{Test}

$$J_{\text{Train}} = \sum_{i \in \text{Train set}} \|d_i - y_i\|^2 \quad \text{και} \quad J_{\text{Test}} = \sum_{i \in \text{Test set}} \|d_i - y_i\|^2$$

όπου Train Set, Test Set είναι το σύνολα των δεδομένων εκπαίδευσης και ελέγχου, ενώ d_i , y_i είναι αντίστοιχα οι στόχοι και έξοδοι του δικτύου για το πρότυπο t_i .



Όταν η πολυπλοκότητα του δικτύου ξεπεράσει μια κρίσιμη τιμή, η ικανότητα γενίκευσης μειώνεται, καθώς το J_{Test} αυξάνει, παρά το γεγονός ότι το σφάλμα εκπαίδευσης J_{Train} εξακολουθεί να μειώνεται



Το ζητούμενο είναι να βρεθεί μια διαχωριστική γραμμή μεταξύ των πορτοκαλί και των μπλε προτύπων

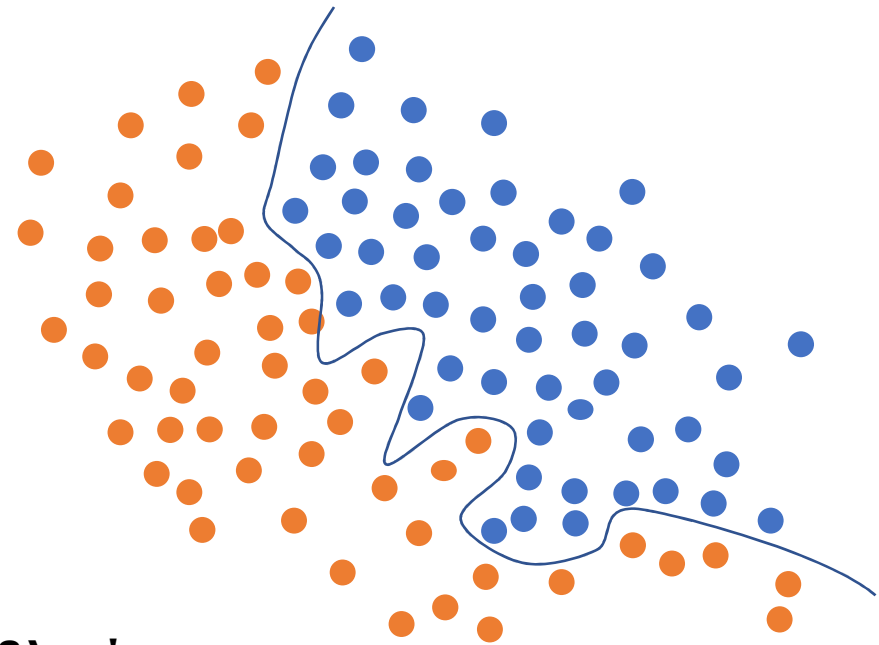
Υπέρ-Εκπαίδευση ή Υπερ-προσαρμοσμένη (Overfitting)

Το δίκτυο μοντελοποιεί ανεπιθύμητες λεπτομέρειες (θόρυβο) μεταξύ των δύο κατηγοριών.

Μικρό σφάλμα γενίκευσης, όμως το δίκτυο αδυνατεί να αποδώσει αντίστοιχα σε άγνωστα πρότυπα

Χαρακτηριστικές Μέθοδοι Μετριάσμού των Προβλημάτων:

- Φθορά των Βαρών ή κανονικοποίηση (Regularization)
- Πρόωρο Σταμάτημα (Early Stopping)



Το ζητούμενο είναι βρεθεί μια διαχωριστική γραμμή μεταξύ των πορτοκαλί και των μπλε προτύπων

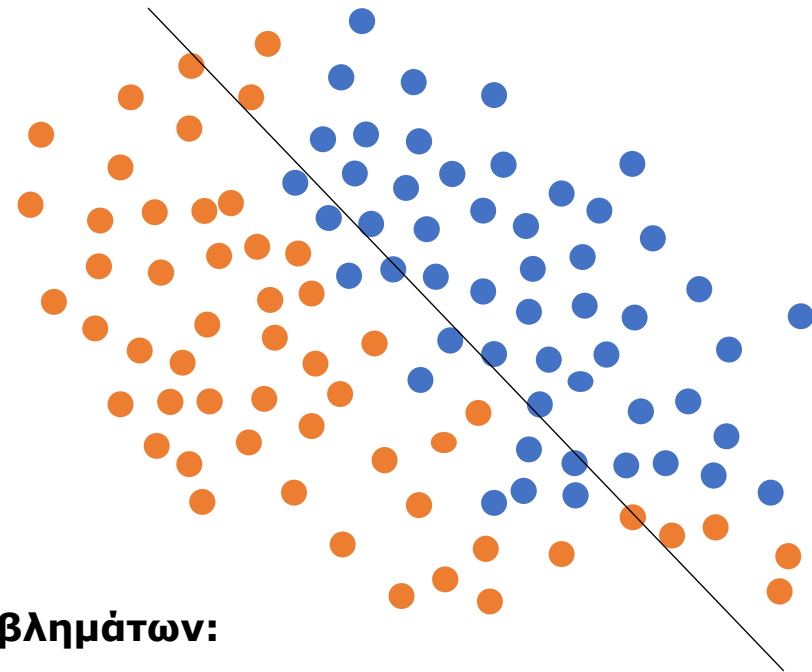
Υπό-Εκπαίδευση ή Υπό-Προσαρμοσμένη ή Ατελής Μάθηση (Underfitting):

Το δίκτυο είναι πολύ απλοϊκό για να μοντελοποιήσει ικανοποιητικά τη διαχωριστική γραμμή μεταξύ των δύο κατηγοριών

Μεγάλο σφάλμα γενίκευσης. Το δίκτυο αδυνατεί να συγκλίνει στις επιθυμητές εξόδους

Χαρακτηριστικές Μέθοδοι Μετριάσμού των Προβλημάτων:

- Φθορά των Βαρών ή κανονικοποίηση (Regularization)
- Πρόωρο Σταμάτημα (Early Stopping)



Εργαλεία Λογισμικού



Εργαλεία Λογισμικού

- WEKA
- RapiMiner
- Matlab Neural Network Toolbox
- Wolfram Mathematica Neural Networks
- R nnet, neuralnet, RSNN
- Theano
- Microsoft Cognitive Toolkit
- Keras
- Apache MXNet
- SPSS Clementine
- SAS Enterprise Miner
- PredictorPro
- NeuroSolutions
- Alyuda NeuroSolutions
- Neural Power
- NeurShell Predictor
- NeuroLab
- Neural Works Professional
- JNNS JvaSNNS
- SNNS

