

Τεχνητή νοημοσύνη νευρώνες ή συναρτήσεις;

Δρ.Ηλ.Μηχ. Ερρίκος Βεντούρας

Τμήμα Μηχανικών Βιοϊατρικής

Πανεπιστήμιο Δυτικής Αττικής

Στοιχείο της νοημοσύνης είναι η αντίληψη ότι:

- Σε μια **εικόνα** υπάρχουν (ή δεν υπάρχουν) συγκεκριμένα όντα και αντικείμενα. Υπάρχει π.χ. στη εικόνα μία γάτα του Σιάμ;
- Σε ένα **ήχο** υπάρχει ομιλία ή μουσική.
 - *Ποιες είναι οι προτάσεις που εκφέρονται στην ομιλία;*
 - *Ποιο το μουσικό κομμάτι που ακούμε;*

Στοιχεία της νοημοσύνης:

Όταν μας ζητείται να γράψουμε μια αναφορά για ένα ιστορικό γεγονός πρέπει:

- να αντιληφθούμε/σκεφθούμε **ποιο είναι αυτό το γεγονός** και να ανταποκριθούμε με επιτυχία στα παρακάτω **ερωτήματα**:
- ❖ Σε **ποια βιβλία** μιας βιβλιοθήκης μπορεί να βρούμε σχετικές πληροφορίες, δηλ. **κείμενο**;
- ❖ Όταν ανοίξουμε τα βιβλία, **σε ποια κεφάλαια** πιθανότατα υπάρχει το σχετικό κείμενο;
- ❖ Ποια **τμήματα του κειμένου** να διαλέξουμε για να διαβάσουμε;
- ❖ Αφού συλλέξουμε και διαβάσουμε κείμενα, πώς να τα **ανασυνθέσουμε**;

Όλα τα παραπάνω στηρίζονται σε μεγάλο βαθμό στην ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ

- Έστω ότι διακρίνουμε 2 ξένα σύνολα

(ή ισοδύναμα κλάσεις, καταστάσεις, τάξεις):

- *Τμήματα εικόνων που περιέχουν γάτες (κλάση c_1)*
 - *Τμήματα εικόνων που ΔΕΝ περιέχουν γάτες (κλάση c_2)*
- Έστω ότι για κάθε τμήμα εικόνας μετράμε **N χαρακτηριστικά**
(αριθμούς, πίνακες)

*π.χ. καμπύλωση αντίθεσης και συσχέτισή της με γειτονικά τμήματα, τόνοι
χρωμάτων κ.ά.*

Το θεμελιώδες ερώτημα στην αναγνώριση των προτύπων είναι το εξής:

Όταν «εξετάσουμε» ένα τμήμα εικόνας
για το οποίο δεν γνωρίζουμε εκ των προτέρων σε ποια κλάση ανήκει
είναι δυνατόν "διαβάζοντας την υπογραφή του"
δηλ. εξετάζοντας το διάνυσμα χαρακτηριστικών του
να αποφασίσουμε σε ποια κλάση ανήκει;



Υπάρχει κάποιος κατάλληλος κανόνας (ταξινομητής)
βάσει του οποίου

με μόνο διαθέσιμο στοιχείο το εκάστοτε διάνυσμα \underline{f}_q
να μπορούμε να ταξινομούμε το \underline{f}_q σε μία από τις υπάρχουσες κλάσεις;



Και αν υπάρχει τέτοιος ταξινομητής
είναι ο βέλτιστος για την περίπτωσή μας;
δηλ. μας οδηγεί σε απόφαση ταξινόμησης
με το ελάχιστο δυνατό ποσοστό σφάλματος;
Το κριτήριο απόφασης, δηλ. ο ταξινομητής
μπορεί να στηρίζεται σε μία συνάρτηση ταξινόμησης g
με είσοδο το διάνυσμα χαρακτηριστικών και έξοδο μία τιμή (0 ως 1)

ΚΑΝΟΝΑΣ ΤΑΞΙΝΟΜΗΣΗΣ:

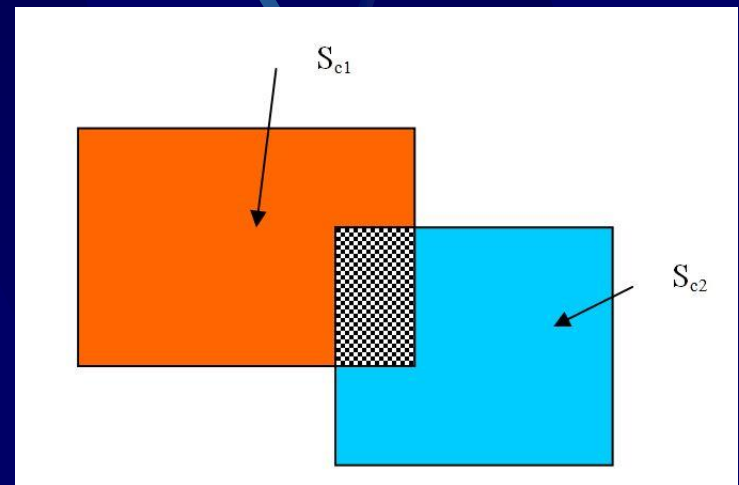
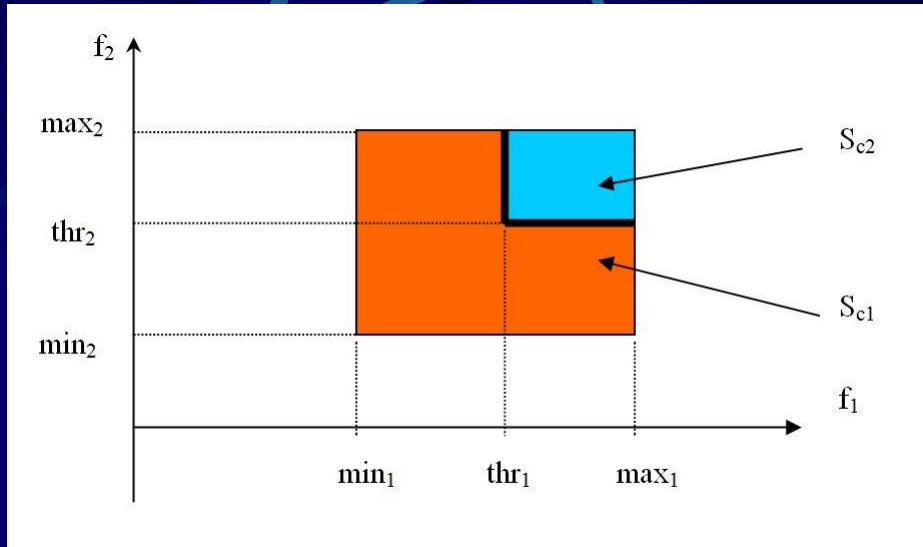
Εστω κλάσεις $j=1, \dots, K$ και κάθε κλάση έχει συνάρτηση ταξινόμησης g_j

Τότε ταξινόμησε το διάνυσμα f_q

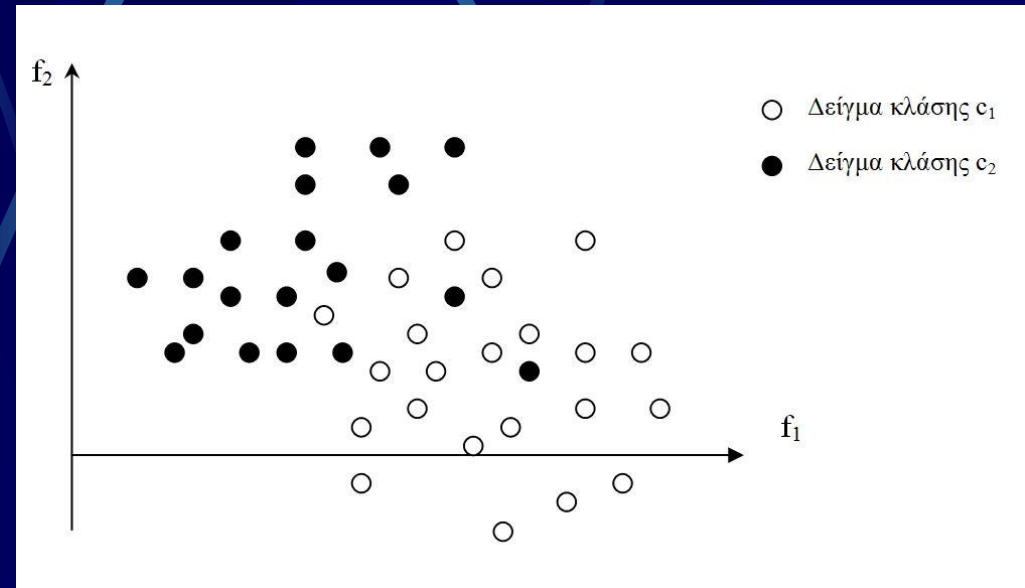
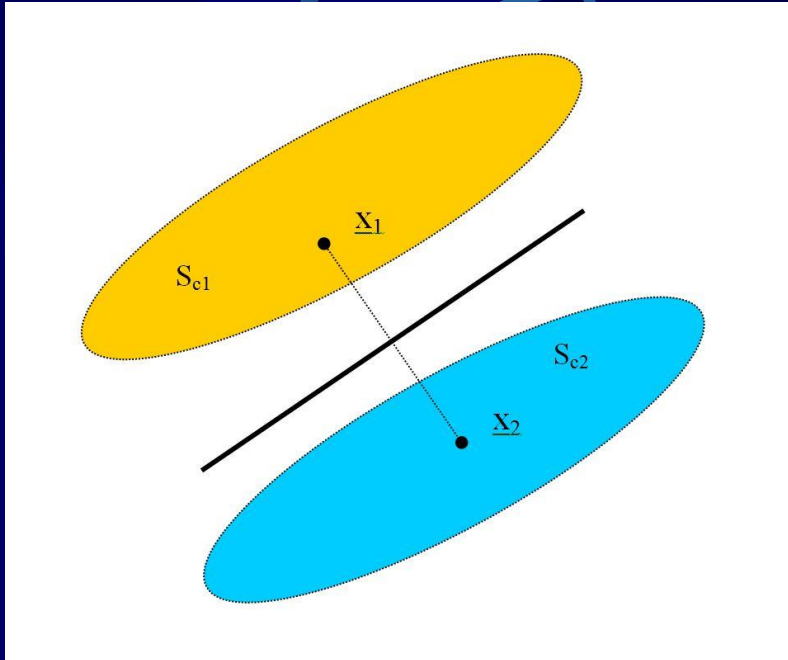
(δηλ. ταξινόμησε το τμήμα εικόνας q από το οποίο προήλθε το διάνυσμα f_q) στην κλάση X

όταν $g_X(f_q) > g_j(f_q)$ για κάθε $j=1, \dots, K$ με $j \neq X$

Χώρος των χαρακτηριστικών – περιοχές των κλάσεων



Χώρος των χαρακτηριστικών – περιοχές των κλάσεων



Οι διάφοροι «κλασικοί» «ταξινομητές», «συναρτήσεις απόφασης», π.χ. κ πλησιεστέρων γειτόνων, γκαουσιανός ταξινομητής κ.ά. έχουν **σχετικά καλή απόδοση** σε προβλήματα με γεωμετρικώς «εύκολα» διαχωρίσιμες περιοχές των κλάσεων.

Ο ανθρώπινος εγκέφαλος είναι γνωστό ότι έχει **εκπληκτικές ικανότητες εκμάθησης προτύπων** και **ανίχνευσης των προτύπων αυτών σε νέα δεδομένα.**

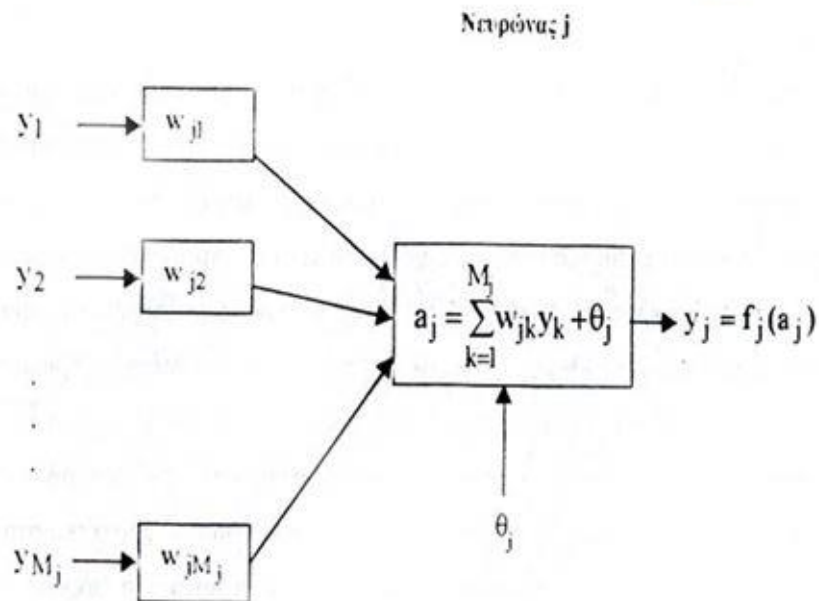
Οι γνωσιακές λειτουργίες του εγκεφάλου υλοποιούνται από τους **νευρώνες.**

Μπορούν μαθηματικά πρότυπα της λειτουργίας των νευρώνων να αποτελέσουν ταξινομητές;



Δομή του νευρώνα
με το σώμα, τον
άξονα και τους
δενδρίτες

ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΕΝΟΣ ΤΕΧΝΗΤΟΥ ΝΕΥΡΩΝΑ (j)

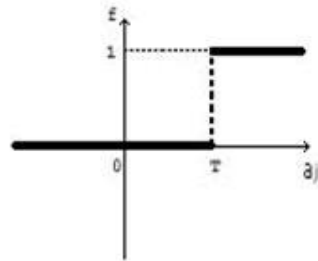


όπου:

- j ο αύξων αριθμός των νευρώνων του δικτύου: $1, 2, \dots, N$
- k ο αύξων αριθμός των συνάψεων του κάθε νευρώνα με τους «προηγούμενους»: $1, 2, \dots, M_j$
- y_k το σήμα εξόδου του k νευρώνα που τροφοδοτεί τον j νευρώνα
- w_{jk} το βάρος της σύναψης από τον k νευρώνα στον j νευρώνα
- a_j κατάσταση ενεργοποίησης του νευρώνα j
- θ_j παράμετρος πόλωσης

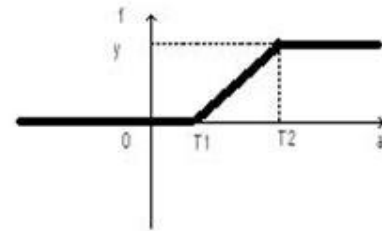
Συναρτήσεις εισόδου-εξόδου f του Τεχνητού Νευρώνα

(α)



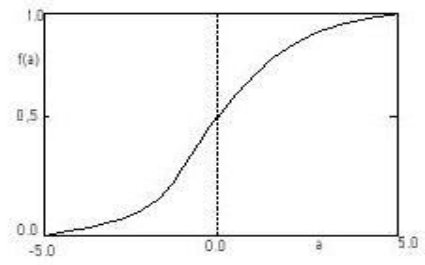
βηματική

(β)



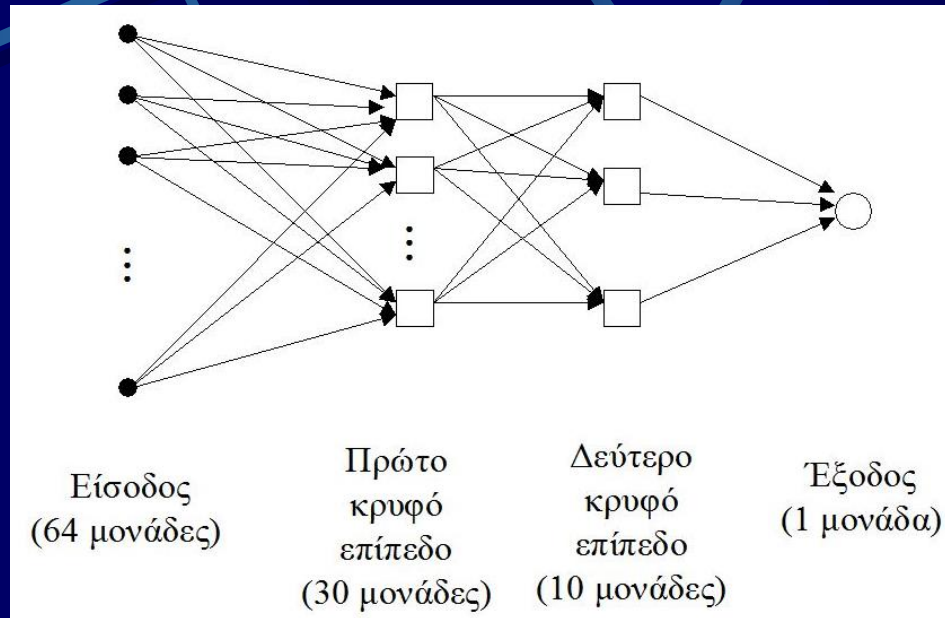
γραμμική

(γ)



σιγμοειδής

Τεχνητό Νευρωνικό Δίκτυο



Στοιχεία του ΤΝΔ:

Ο αριθμός των εισόδων, ο αριθμός των στρωμάτων, ο αριθμός των νευρώνων σε κάθε στρώμα, η συνάρτηση ενεργοποίησης, τα συναπτικά βάρη.

Το νευρωνικό δίκτυο, όπως και κάθε μεμονωμένος νευρώνας, υλοποιεί εντέλει μια συνάρτηση.

Τεχνητό Νευρωνικό Δίκτυο (ΤΝΔ)

- Η κύρια ιδιότητα του εγκεφάλου που μας ενδιαφέρει είναι η **εκμάθηση**. Συγκεκριμένα, δίνοντας μας εικόνες από γάτες, αλλά και εικόνες από άλλα ζώα, **οι νευρώνες μας θα «διαπλαστούν»** έτσι ώστε κάποια στιγμή να ξεχωρίζουμε τις γάτες από άλλα ζώα όταν τις δούμε.
- Είναι γνωστό ότι η **συναπτική πλαστικότητα** είναι σημαντικός παράγοντας σε διαδικασίες μάθησης στον εγκέφαλο.
- Σε ένα ΤΝΔ **είσοδοι είναι τα διανύσματα χαρακτηριστικών** και **έξοδος η απόφαση ένταξης σε κλάση**.
- Αν έχουμε αποφασίσει την αρχιτεκτονική ενός δικτύου, αυτό που θα καθορίσει εντέλει την **συνάρτηση εισόδου-εξόδου** του ΤΝΔ είναι **οι τιμές των συναπτικών βαρών**.
- *Πώς μπορούμε να υπολογίσουμε τα βάρη, χρησιμοποιώντας διανύσματα χαρακτηριστικών από δείγματα από τις κλάσεις που μας ενδιαφέρουν;*

ΑΛΓΟΡΙΘΜΟΣ ΕΚΜΑΘΗΣΗΣ BACK PROPAGATION

Βήμα 1:

Τα βάρη $w_{jk}^{L,1}$ και οι παράμετροι πόλωσης θ_j^L λαμβάνουν μικρές τυχαίες τιμές.

Βήμα 2:

Παρουσιάζεται κάποιο διάνυσμα εισόδου $\underline{x}^{(1)} = (x_1^{(1)}, \dots, x_n^{(1)})$, όπου $n=N_1$. Ο εκθέτης (1) δηλώνει ότι αυτό είναι το πρώτο διάνυσμα που παρουσιάζουμε στην διαδικασία της εκμάθησης (δηλ. βρισκόμαστε στο πρώτο στάδιο της εκμάθησης, $Q=1$). Το διάνυσμα εκμάθησης $\underline{x}^{(1)}$ είναι επιλεγμένο από το σύνολο των διανυσμάτων χαρακτηριστικών που διαθέτουμε προς εκμάθηση και έστω ότι ανήκει στην κλάση c_m , η οποία αντιπροσωπεύεται από διάνυσμα $\underline{t}^{(m)} = (t_1^{(m)}, \dots, t_{n'}^{(m)})$, όπου $n'=N_{Lmax}$. Η επιθυμητή δηλαδή έξοδος του ΤΝΔ είναι το $\underline{t}^{(m)}$.

Βήμα 3:

Υπολογίζουμε την έξοδο κάθε νευρώνα j κάθε στρώματος L

Βήμα 4:

Τα βάρη των συνάψεων που συνδέουν δύο γειτονικά στρώματα (L και L-1) αναπροσαρμόζονται, ξεκινώντας από το στρώμα εξόδου και οδεύοντας στο στρώμα εισόδου, δηλ. οι παρακάτω τύποι εφαρμόζονται διαδοχικά για $L = L_{\max}$, $L_{\max}-1$, $L_{\max}-2$, ..., 1:

$$w_{jk}^{L,2} = w_{jk}^{L,1} + \eta \delta_j^{L,1} y_k^{L-1,1} \quad j = 1, \dots, N_L \quad (5.8)$$

όπου η είναι ο ρυθμός μάθησης και κανονίζει τον ρυθμό αναπροσαρμογής των βαρών και $\delta_j^{L,1}$ είναι το σήμα σφάλματος του j νευρώνα του L στρώματος στο πρώτο στάδιο εκμάθησης ($Q=1$).

Αν το σήμα σφάλματος προκύψει μικρότερο από ένα προκαθορισμένο όριο τότε θεωρείται μηδενικό.

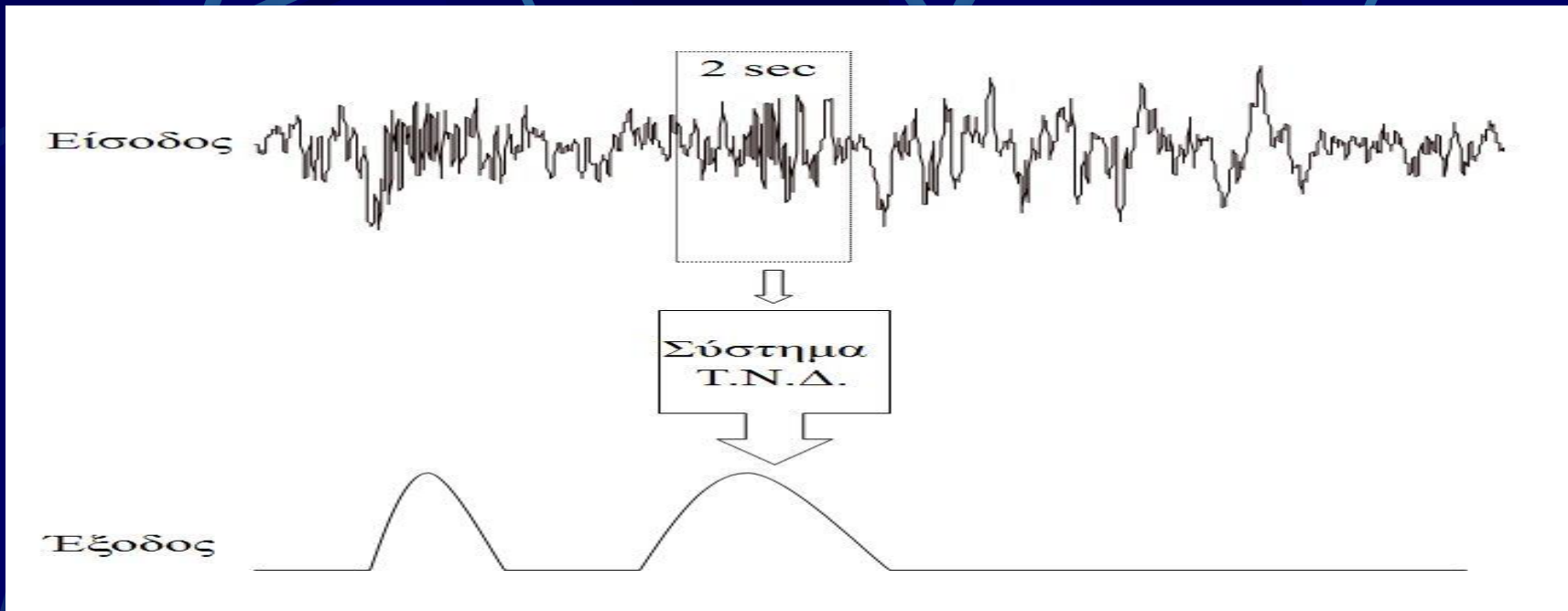
Βήμα 5:

Με βάση τα βάρη που υπολογίστηκαν στο Βήμα 4, παρουσιάζουμε στην είσοδο του ΤΝΔ το επόμενο διάνυσμα εκμάθησης και για $Q=2$ επαναλαμβάνουμε τα βήματα 2 ως και 4. Το ίδιο γίνεται στην συνέχεια για όλα τα διανύσματα εκμάθησης που διαθέτουμε, συμπληρώνοντας δηλαδή ένα πλήρη κύκλο εκμάθησης.

Βήμα 6:

Αν στο τέλος του κύκλου διαπιστώσουμε ότι τα βάρη δεν άλλαξαν καθόλου τότε η εκμάθηση θεωρείται επιτυχής, αλλιώς επαναλαμβάνουμε τα βήματα από το 2 και ύστερα.

Παράδειγμα εισόδου σε ΤΝΔ: κυματομορφή εξελισσόμενη στον χρόνο



Παράδειγμα εξόδου σε ΤΝΔ: η τιμή εξόδου ενός νευρώνα. Ανιχνεύεται ένα συγκεκριμένο, ως προς τις συχνότητες, την μορφολογία και το πλάτος, είδος κυματομορφής.

«Βαθειά» μάθηση: πολυστρωματικά δίκτυα

- Όταν ο χώρος των χαρακτηριστικών γίνεται πολύπλοκος τα ΤΝΔ με **1 έως 3 «κρυφά»/ενδιάμεσα στρώματα** δεν επιτυγχάνουν ικανοποιητική ταξινόμηση.
- Η χρήση **πολλών ενδιάμεσων στρωμάτων** επιφέρει πολύ μεγάλες υπολογιστικές απαιτήσεις και ο «κλασικός» αλγόριθμος “back-propagation” καθίσταται ανεπαρκής.

Τα τελευταία 15 χρόνια έγινε εφικτή η λεγόμενη **«βαθειά» μάθηση και τα «βαθειά» δίκτυα**, χάρη:

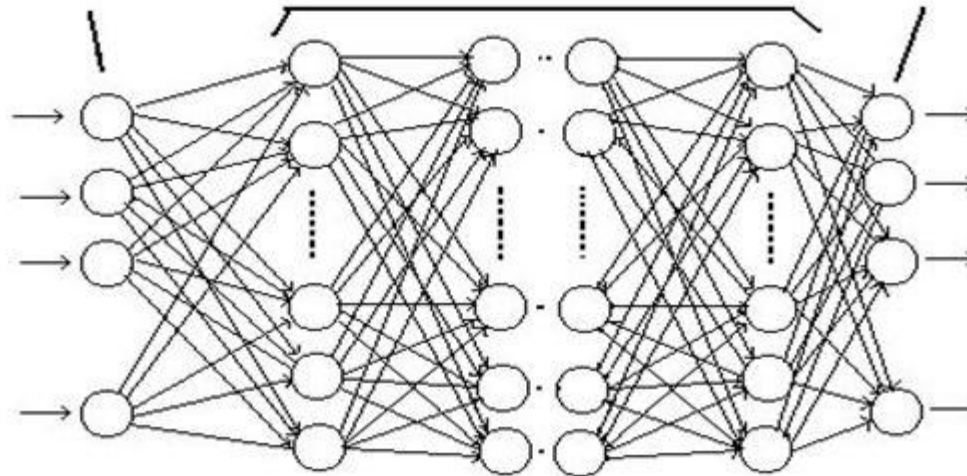
- στην κατακόρυφη αύξηση της υπολογιστικής ισχύος, ιδίως με εξειδικευμένες κάρτες γραφικών &
- στην ανάπτυξη νέων αλγορίθμων εκμάθησης.

Τοπολογίες ευθείας τροφοδότησης με πολλά ενδιάμεσα/κρυφά στρώματα

Στρώμα Εισόδου

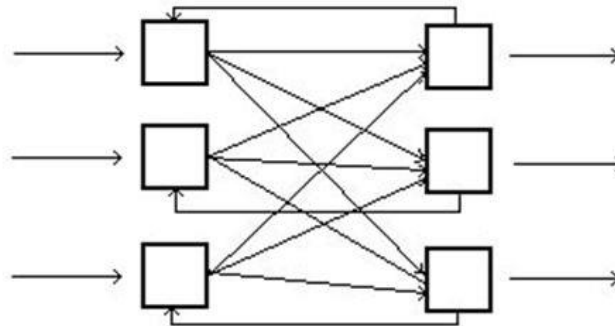
Κρυφά στρώματα

Στρώμα Εξόδου



- ❖ Ένα πολύ ενδιαφέρον στοιχείο στην ανθρώπινη μάθηση είναι η ικανότητά μας, όταν π.χ. ακούσουμε ένα γνώριμο μουσικό σκοπό, ξεκινώντας από ένα μικρό μέρος της μελωδίας, να μπορούμε να **αναπαράγουμε στο μυαλό μας και να τραγουδήσουμε την συνέχεια του μουσικού κομματιού**.
- Τα ΤΝΔ «ευθείας» τροφοδότησης, δηλ. με «μονόδρομες» συναρτήσεις από την είσοδο προς την έξοδο, δεν ανταποκρίνονται σε τέτοιες απαιτήσεις.
- ✓ Η λύση βρίσκεται στην χρήση δικτύων με **ανάδραση**.

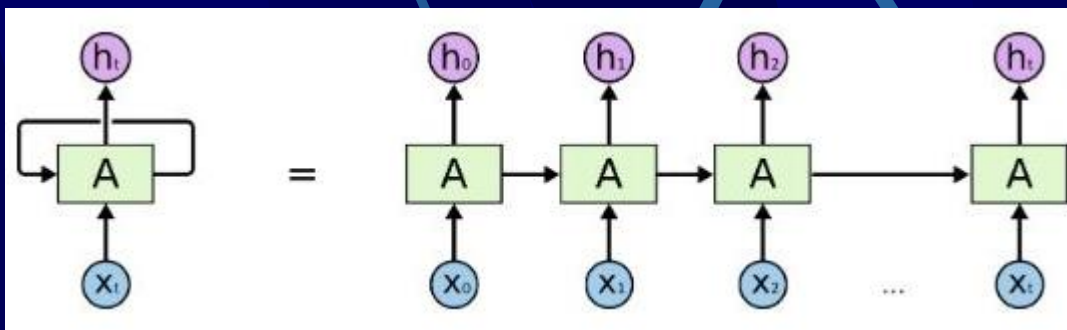
Εναλλακτική αρχιτεκτονική:
Δίκτυο δυο στρωμάτων με ανάδραση



Ανάδραση: άμεσα συνδεδεμένη με χρονικά αλληλοδιάδοχες και συσχετισμένες μνήμες.

Αριστερά: συνεπτυγμένη σχηματική αναπαράσταση δικτύου με ανάδραση.

Δεξιά: ανεπτυγμένη σχηματική αναπαράσταση. Δείχνεται ότι η επόμενη κατάσταση του δικτύου εξαρτάται από την προηγούμενη (τουλάχιστον) κατάσταση.



Παράδειγμα:

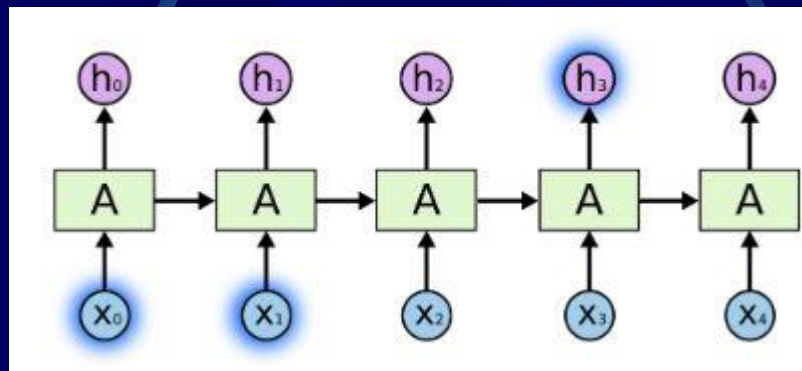
Η ανάκληση/επιλογή μιας «επόμενης» μνήμης/λέξης

σχετίζεται με αμέσως

προηγούμενες: «τα **σύννεφα**

(X0) εμφανίζονται (X1) πάντα

στον >>> ουρανό (h3)»



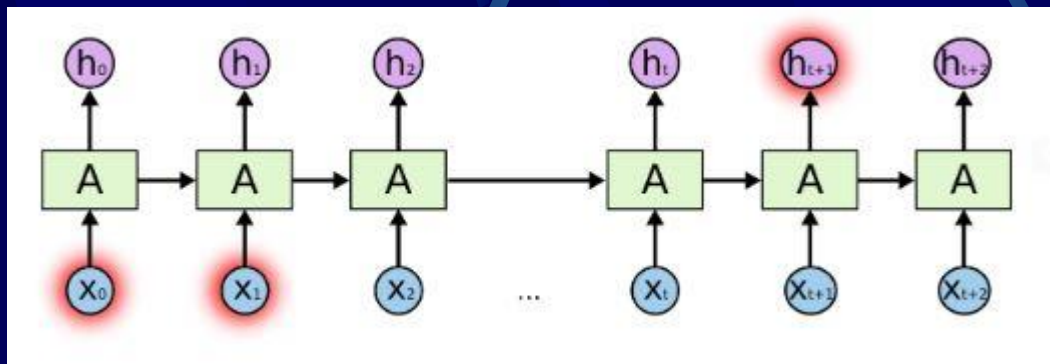
Υπάρχουν περιπτώσεις που χρειαζόμαστε πληροφορία από αρκετά «παλαιότερη» κατάσταση:

Παράδειγμα:

Η ανάκληση/επιλογή μιας «επόμενης» μνήμης/λέξης σχετίζεται με λέξη/πληροφορία που εμφανίστηκε **αρκετά προτούτερα σε ένα κείμενο**.

«Έχω **γεννηθεί** στην **Γαλλία** (X_0), μεγάλωσα και πήγα **σχολείο** στην **Ρουέν** (X_1), ..., η γλώσσα που μιλώ καλύτερα είναι βεβαίως τα **γαλλικά** ($h(t+1)$).

Η αρχική πληροφορία που θα καθορίσει την σωστή απάντηση μπορεί να βρίσκεται πολλές προτάσεις «νωρίτερα» στο κείμενο.

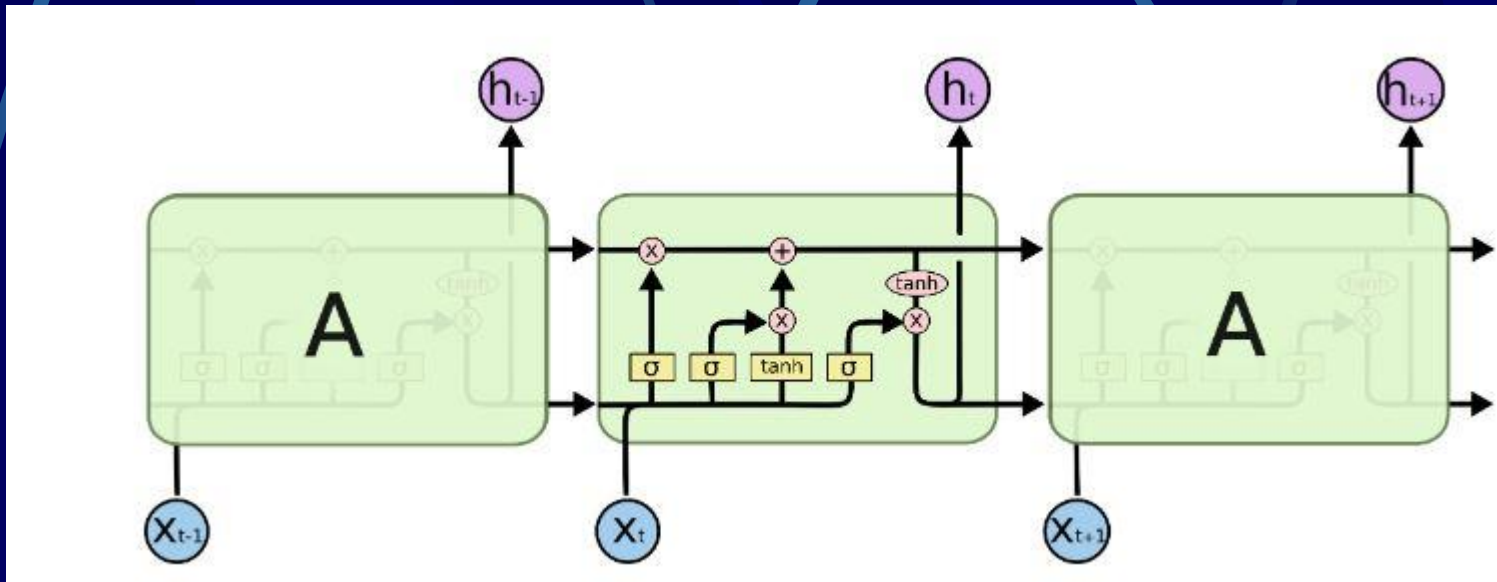


- Σημειώστε ότι στα προηγούμενα ΔΕΝ χρησιμοποιούμε γραμματικούς ή συντακτικούς κανόνες ή τυπική λογική για να παράξουμε την κατάλληλη επόμενη πρόταση.
- Στηριζόμαστε στην **συχνότερα εμφανιζόμενη αλληλουχία λέξεων και προτάσεων** σε αυτό ή σε άλλα κείμενα.
- Χρειάστηκε να βρεθεί ένα είδος δικτύου το οποίο «μαθαίνει» και «θυμάται» μνήμες **«βραχείες» και «μακρές»:**

**Δίκτυο μακρο-βραχυπρόθεσμης μνήμης
Long-Short Term Memory (LSTM) network**

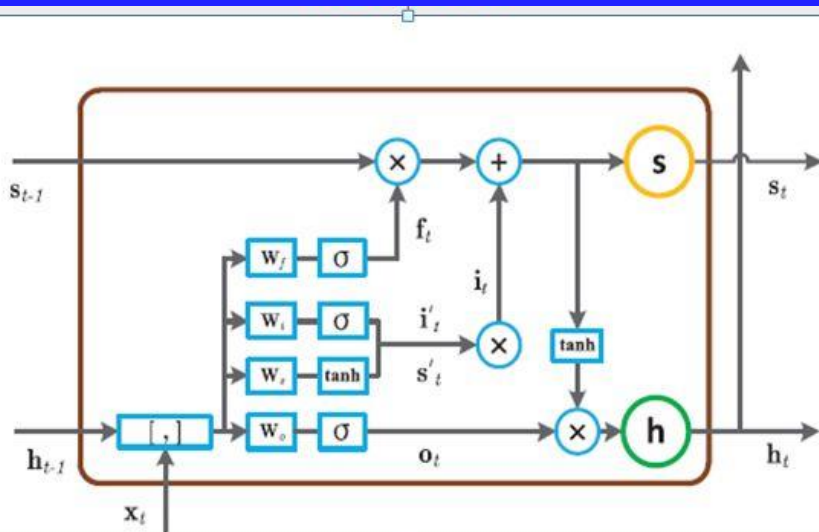
Δίκτυο μακρο-βραχυπρόθεσμης μνήμης Long-Short Term Memory (LSTM) network:

- ❖ Το δίκτυο (A) περιέχει πολύπλοκη εσωτερική δομή.
- ❖ Στο παρακάτω σχήμα δείχνεται δίκτυο με μια «μονάδα» LSTM. Κάθε ορθογώνιο στοιχείο περιέχει αλληλοσυνδεόμενα επιμέρους ΤΝΔ. Στο σχήμα αναγράφεται το είδος της συνάρτησης ενεργοποίησης κάθε επιμέρους ΤΝΔ.
- ❖ Το δίκτυο (A) μπορεί να περιέχει πολλές τέτοιες «μονάδες», αλληλοσυνδεόμενες.



Δίκτυο μακρο-βραχυπρόθεσμης μνήμης Long-Short Term Memory (LSTM) network:

Αναλυτική δομή μιας μονάδας LSTM



Συναρτήσεις

- $f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$ (**forget gate**)
- $i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$ (**input gate**)
- $\tilde{s}_t = \tanh(W_s \cdot [h_{t-1}, x_t] + b_s)$
- $s_t = f_t * s_{t-1} + i_t * \tilde{s}_t$ (**state**)
- $o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o)$ (**output gate**)
- $h_t = o_t * \tanh(s_t)$ (**hidden state**)

Άλλες προσεγγίσεις: ΤΝΔ χωρικής συσχέτισης

Δίκτυα για **αναγνώριση εικόνων-μορφών**

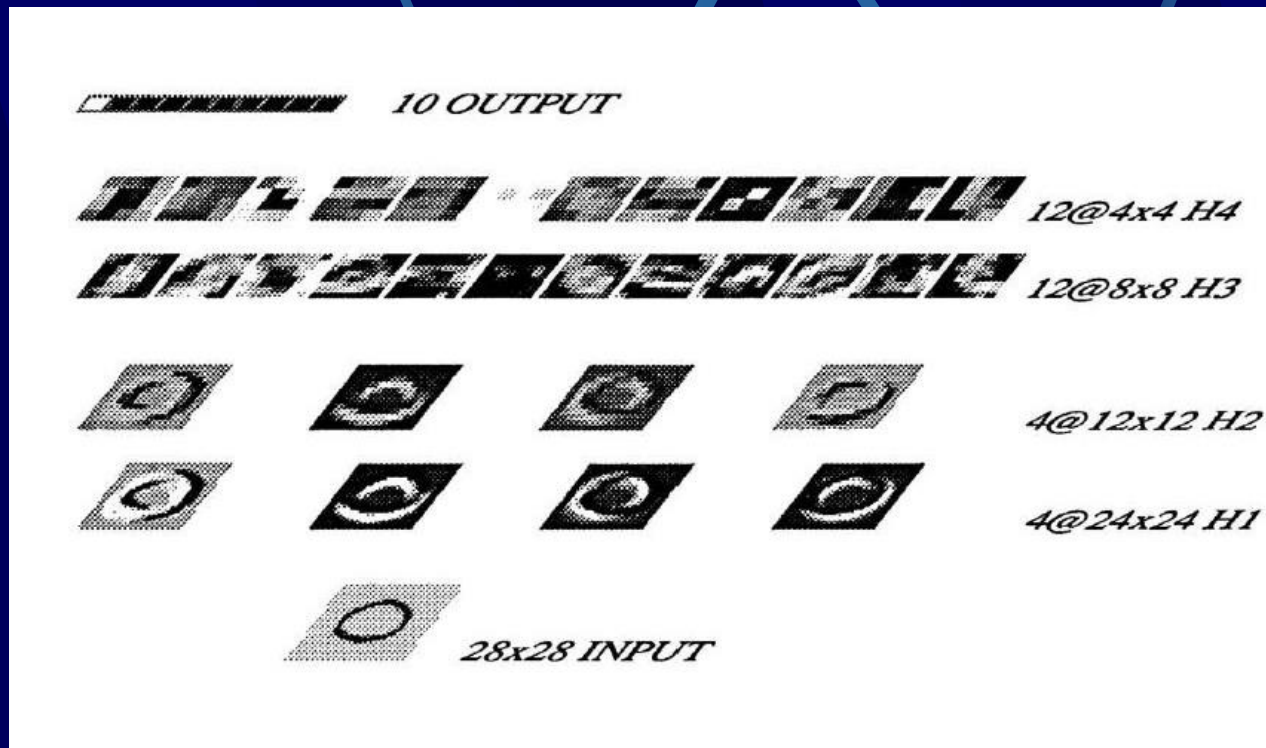
Η βασική ιδέα είναι να χωρίζουμε την εικόνα σε τμήματα και να υπολογίζουμε την **συνέλιξη μεταξύ των εικονοστοιχείων**.

Στη συνέχεια επαναλαμβάνεται η διαδικασία εξαγωγής συνέλιξης στα επόμενα επίπεδα, έως το επίπεδο εξόδου το οποίο επιλέγει/προτείνει μεταξύ σχημάτων/μορφών (κλάσεις).

Τα δίκτυα αυτά ονομάζονται **νευρωνικά δίκτυα συνέλιξης – convolutional neural networks (CNN)**.

Άλλες προσεγγίσεις: ΤΝΔ χωρικής συσχέτισης

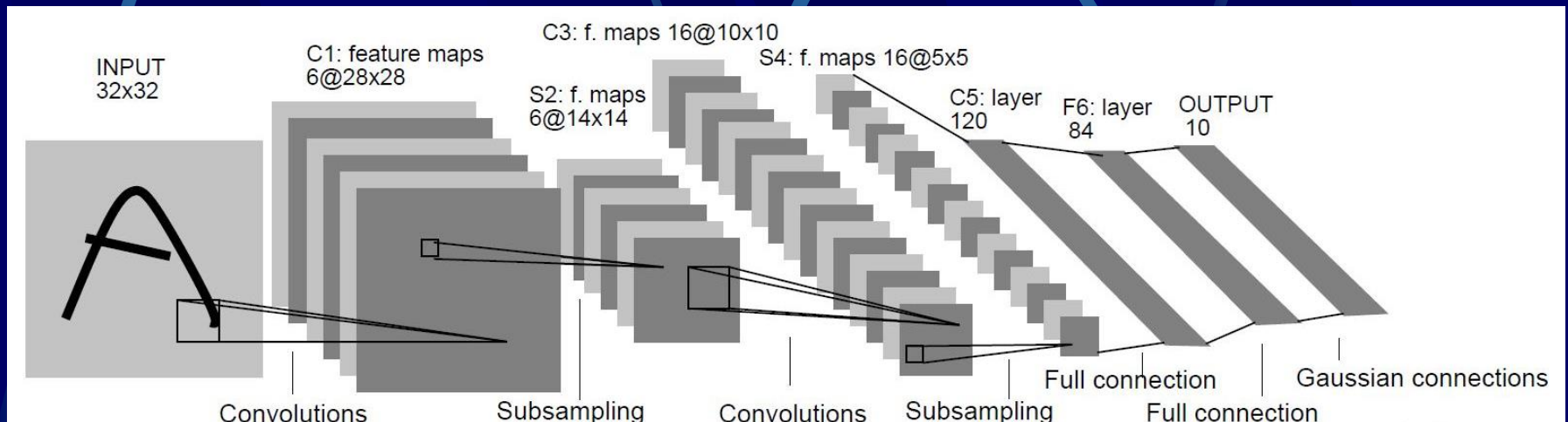
Η έρευνα ξεκίνησε την δεκαετία του 1980, για να λυθεί το πρόβλημα της αυτόματης ανάγνωσης ταχυδρομικών φακέλων.



Άλλες προσεγγίσεις: ΤΝΔ χωρικής συσχέτισης

Με την έλευση της **βαθιάς μάθησης** μπόρεσαν να χρησιμοποιηθούν σε όλο
δυσκολότερες εργασίες:

όχι πλέον μόνο οι 20+ χαρακτήρες και τα δέκα ψηφία,
αλλά **εκατομμύρια μορφές σε εικόνες και βίντεο**.



Το μέλλον: συνδυασμοί δικτύων, αύξηση της υπολογιστικής δύναμης, νέοι αλγόριθμοι εκμάθησης ...

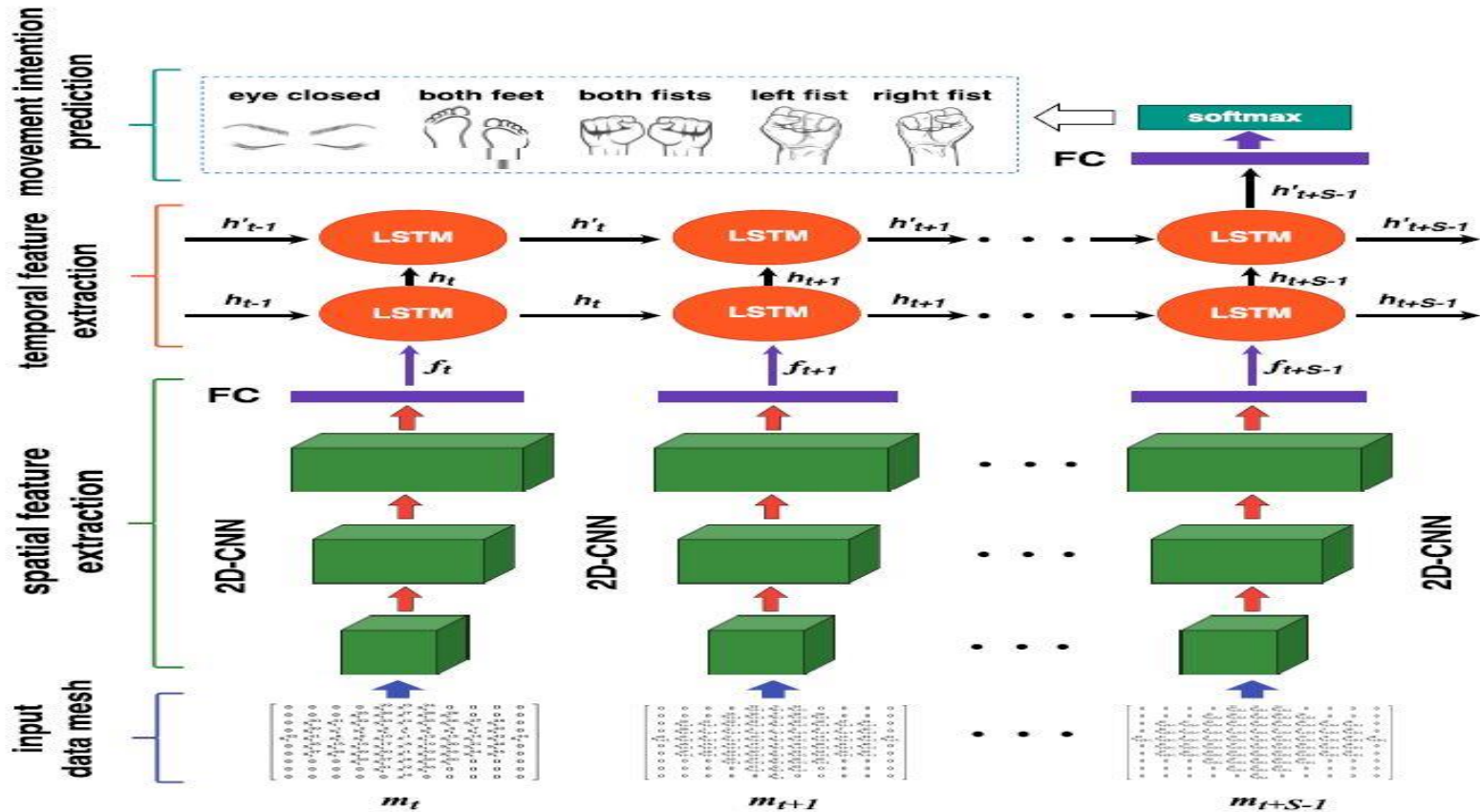


Fig. 2. Cascade convolutional recurrent neural network architecture.

Ευχαριστώ!