

1. ΕΙΣΑΓΩΓΗ

1.1 ΤΙ ΕΙΝΑΙ ΕΝΑ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ;

Η εργασία σε τεχνητά νευρωνικά δίκτυα, κοινώς γνωστά ως “νευρωνικά δίκτυα”, υποκινήθηκε άμεσα από την αναγνώριση ότι οι εγκεφαλικοί υπολογιστές είναι ένας εξ’ ολοκλήρου διαφορετικός δρόμος από τους συμβατικούς ψηφιακούς υπολογιστές. Τυπικά, οι νεύρωνες είναι 5-6 τάξεις μεγέθους αργότεροι από τις λογικές πύλες σιλικόνης.

Ο εγκέφαλος είναι ένας πολύ πολύπλοκος, μη-γραμμικός και παράλληλος υπολογιστής. Έχει τη δυνατότητα να οργανώνει τους νεύρωνες έτσι ώστε να εκτελεί συγκεκριμένους υπολογισμούς πολύ πιο γρήγορα από τους πιο γρήγορους ψηφιακούς υπολογιστές που υπάρχουν.

Πώς τα καταφέρνει όμως ο εγκέφαλος; Κατά τη γέννησή του ο εγκέφαλος έχει την ικανότητα να κατασκευάζει τους δικούς του κανόνες, κοινώς “εμπειρία”, η οποία μεγαλώνει με τα χρόνια. Κατά τα 2 πρώτα χρόνια ζωής, έχουμε τη μέγιστη ανάπτυξη, όπου περίπου 1 εκατομμύριο συνάψεις (synapses) δημιουργούνται στο δευτερόλεπτο.

Οι synapses είναι οι βασικές δομικές και λειτουργικές μονάδες που μεσολαβούν στην ενδοεπικοινωνία των νευρώνων.

Άρα:

Ένα νευρωνικό δίκτυο είναι ένα συμπαγές παράλληλος καταναμημένος επεξεργαστής, που έχει τη φυσική κλίση να αποθηκεύει εμπειριστατωμένη γνώση και να την κάνει διαθέσιμη για χρήση.

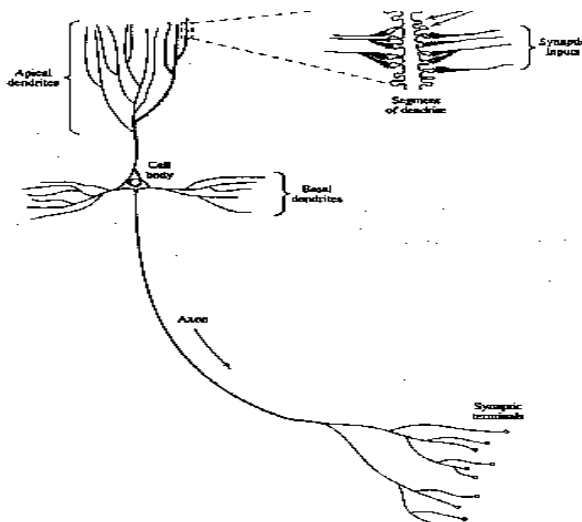


FIGURE 1.1 The pyramidal cell.

Μοιάζει με τον εγκέφαλο στα εξής:

- Η γνώση αποκτάται από το δίκτυο μέσα από διαδικασία εκμάθησης.

•Οι δυνάμεις σύνδεσης των νευρώνων, γνωστές σαν συναπτικά (synaptic) βάρη, χρησιμοποιούνται για την αποθήκευση γνώσης.

Η διαδικασία για τη εκμάθηση ονομάζεται “αλγόριθμος εκμάθησης”. Η μεταβολή των συναπτικών (synaptic) βαρών, αποτελεί την παραδοσιακή μέθοδο για το σχεδιασμό των νευρωνικών δικτύων.

Πλεονεκτήματα των Νευρωνικών Δικτύων

Η συμπαγής παράλληλα κατανεμημένη δομή και η ικανότητα μάθησης των νευρωνικών δικτύων, κάνουν δυνατή την επίλυση πολύπλοκων προβλημάτων, τα οποία σπάνε σε μικρότερα έργα, που αναλαμβάνουν τα νευρωνικά δίκτυα ανάλογα με τις δυνατότητές τους.

Η χρήση των Νευρωνικών Δικτύων παρέχει τις ακόλουθες χρήσιμες ιδιότητες :

1. Μη-γραμμικότητα. Αυτή οφείλεται στο γεγονός ότι ένα Νευρωνικό Δίκτυο δομείται από τη σύνδεση νευρώνων, οι οποίοι είναι μη-γραμμικές συσκευές. Η μη-γραμμικότητα είναι πολύ σημαντική ιδιότητα, ιδιαίτερα αν ο φυσικός μηχανισμός για την παραγωγή των σημάτων εισόδου είναι μη-γραμμικός.

2. Σχεδιασμός Εισόδου-Εξόδου. Ένα συνηθισμένο παράδειγμα μάθησης που καλείται επιβλεπόμενη μάθηση, εμπλέκει μεταβολή των synaptic βαρών του Νευρωνικού Δικτύου, εφαρμόζοντας ένα σύνολο δειγμάτων εξάσκησης ή παραδείγματα έργων. Κάθε παράδειγμα αποτελείται από ένα μοναδικό σήμα εισόδου και την επιθυμητή απόκριση. Η εξάσκηση του δικτύου επαναλαμβάνεται για πολλά παραδείγματα, μέχρι το δίκτυο να φτάσει σε μια σταθερή κατάσταση, όπου πλέον δεν γίνονται αλλαγές στα βάρη. Έτσι το δίκτυο μαθαίνει από τα παραδείγματα, κατασκευάζοντας ένα σχεδιασμό εισόδου-εξόδου για το πρόβλημα στο χέρι.

3. Προσαρμοστικότητα. Τα Νευρωνικά Δίκτυα έχουν τη δυνατότητα να προσαρμόζουν τα βάρη τους στις αλλαγές του περιβάλλοντός τους. Μερικές φορές οι προσαρμογές οδηγούν σε μείωση της απόδοσης του συστήματος, γι'αυτό θα πρέπει να είναι επαρκείς δίλημμα σταθερότητας-πλαστικότητας.

4. Αποδεικτική Απόκριση. Ένα Νευρωνικό Δίκτυο σχεδιάζεται για να παρέχει πληροφορίες όχι μόνο για το συγκεκριμένο υπόδειγμα που επιλέγεται αλλά και για την εμπιστοσύνη στην απόφαση που παίρνεται. Αυτό βελτιώνει την απόδοση του συστήματος.

5. Συναφής Πληροφορία. Η γνώση αναπαριστάνεται από την πολύ δομημένη και ενεργή κατάσταση του Νευρωνικού Δικτύου.

6. Αντοχή σε σφάλματα. Ένα νευρωνικό δίκτυο, υλοποιημένο σε hardware φόρμα έχει τη σημαντική ιδιότητα να είναι “ανεκτικό σε σφάλματα”, με την έννοια ότι η απόδοσή του μειώνεται κάτω από αντίξοες λειτουργικές συνθήκες.

7. Υλοποιησιμότητα σε VLSI. Η συμπαγής παράλληλη φύση του Νευρωνικού Δικτύου, κάνει δυνατή την υλοποίηση του σε VLSI τεχνολογία, έτσι ώστε τα νευρωνικά δίκτυα να μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου.

8. Ομοιομορφία Ανάλυσης και Σχεδιασμού. Η έννοια είναι ότι ο ίδιος συμβολισμός χρησιμοποιείται σε όλα τα πεδία που περιέχουν εφαρμογή των νευρωνικών δικτύων. Αυτό το χαρακτηριστικό υποδηλώνεται με διαφορετικούς τρόπους :

Οι νευρώνες παριστάνουν ένα συστατικό συνηθισμένο σε όλα τα νευρωνικά δίκτυα.

Αυτή η συνήθεια κάνει δυνατό το διαμοιρασμό θεωριών και αλγορίθμων εκμάθησης σε διαφορετικές εφαρμογές των νευρωνικών δικτύων.

Ρυθμιστικά δίκτυα μπορούν να κατασκευαστούν μέσω μιας αμιγής ολοκλήρωσης από modules.

9. Αναλογία με Νευροβιολογία. Ο σχεδιασμός νευρωνικών δικτύων γίνεται σε αναλογία με τον εγκέφαλο. Οι νευροβιολόγοι βλέπουν τα νευρωνικά δίκτυα σαν αντικείμενο έρευνας για την εξήγηση νευροβιολογικών φαινομένων. Ομοίως οι μηχανικοί βλέπουν στη νευροβιολογία για νέες ιδέες για την επίλυση πολύπλοκων προβλημάτων.

1.2 ΔΟΜΙΚΑ ΕΠΙΠΕΔΑ ΤΗΣ ΟΡΓΑΝΩΣΗΣ ΤΟΥ ΕΓΚΕΦΑΛΟΥ

Το ανθρώπινο νευρικό σύστημα μπορεί να θεωρηθεί σαν σύστημα 3-επιπέδων. Κεντρικά στο σύστημα είναι ο *εγκέφαλος*, που συνεχώς λαμβάνει πληροφορίες και παίρνει τις κατάλληλες αποφάσεις. Τα *αισθητήρια όργανα* μετατρέπουν τα ερεθίσματα από το ανθρώπινο σώμα ή το εξωτερικό περιβάλλον σε ηλεκτρικούς παλμούς που μεταφέρουν πληροφορία στον εγκέφαλο. Τα *εκτελεστικά όργανα*, αντίθετα, μετατρέπουν τους ηλεκτρικούς παλμούς που γεννιούνται από τον εγκέφαλο σε διακριτές αποκρίσεις στα εξωτερικά συστήματα.

Στον εγκέφαλο υπάρχουν χαμηλής και υψηλής τάξης ανατομικές οργανώσεις και διαφορετικές λειτουργίες. Το σχήμα 1.3 δείχνει μια ιεραρχία των συνυφασμένων επιπέδων οργάνωσης. Ένα *νευρωνικό μικροκύκλωμα*, αναφέρεται σε ένα άθροισμα από *synapses* οργανωμένες σε πρότυπα συνεκτικότητας, ώστε να παρέχουν μια ενδιαφέρουσα λειτουργία. Τα νευρωνικά μικροκυκλώματα ομαδοποιούνται δημιουργώντας *δενδριτικές υπομονάδες* μέσα στα δενδριτικά δέντρα των ξεχωριστών νευρώνων. Στο επόμενο επίπεδο, έχουμε *τοπικά κυκλώματα* που δημιουργούνται από νευρώνες. Ακολουθούν τα *διαπεριφερειακά κυκλώματα* που δημιουργούνται από μονοπάτια, στήλες και τοπογραφικούς χάρτες, που εμπεριέχουν πολλαπλές περιοχές σε διαφορετικά σημεία του εγκεφάλου. Τέλος, οι τοπογραφικοί χάρτες και άλλα διαπεριφερειακά κυκλώματα παρέχουν συγκεκριμένους τύπους συμπεριφοράς στο *κεντρικό νευρικό σύστημα*. Είναι σημαντικό να αναγνωρίσουμε ότι τα δομικά επίπεδα οργάνωσης που περιγράφηκαν είναι μοναδικό χαρακτηριστικό του εγκεφάλου και βρισκόμαστε κοντά στην πραγματοποίησή τους με τεχνητά νευρωνικά δίκτυα.

1.3 ΜΟΝΤΕΛΑ ΝΕΥΡΩΝΑ

Ένας νευρώνας είναι μια μονάδα επεξεργασίας πληροφορίας, που είναι θεμελιακή για την λειτουργία ενός νευρωνικού δικτύου. Το σχήμα 1.4 δείχνει το μοντέλο ενός νευρώνα. Τα 3 βασικά στοιχεία του μοντέλου είναι :

Ένα σύνολο από *synapses* ή *συνδετικούς κρίκους*, καθένα από τα οποία χαρακτηρίζεται από ένα βάρος” ή “δύναμη”. Συγκεκριμένα, ένα σήμα x_j στην είσοδο

της synapse j που συνδέεται στον νεύρωνα K , πολλαπλασιάζεται με το synaptic βάρος w_{kj} . Το βάρος w_{kj} είναι θετικό αν η synapse είναι διεγερτική και αρνητικό αν είναι απαγορευτική.

Έναν *αθροιστή* για την πρόσθεση των σημάτων εισόδου, που παίρνουν βάρος από την αντίστοιχη synapse. Αυτές οι λειτουργίες αποτελούν γραμμικό συνδυαστή.

Μια συνάρτηση ενεργοποίησης για τη μείωση του εύρους της εξόδου του νεύρωνα.

Το μοντέλο επίσης περιλαμβάνει ένα εξωτερικά εφαρμοζόμενο *κατώφλι* θ_k , που έχει επίδραση στην ελάττωση της εισόδου στην επόμενη συνάρτηση ενεργοποίησης. Με μαθηματικούς όρους, ένας νεύρωνα K περιγράφεται με τις εξισώσεις:

$$u_k = \sum_{j=1}^p w_{kj} x_j \quad (1.1)$$

$$y_k = \phi(u_k - \theta_k) \quad (1.2)$$

όπου x_j είναι τα σήματα εισόδου και τα w_{kj} είναι τα synaptic βάρη του K . Το u_k είναι η γραμμική συνδυαστική έξοδος, θ_k το κατώφλι, $\phi(\)$ η συνάρτηση ενεργοποίησης και y_k το σήμα εξόδου του νεύρωνα. Έχουμε όμως μια συγγενική μετατροπή της εξόδου u_k του γραμμικού συνδυαστή από το θ_k ως:

$$v_k = u_k - \theta_k \quad (1.3)$$

Η έξοδος u_k , ανάλογα με το θ_k αν είναι θετικό ή αρνητικό μεταβάλλεται όπως φαίνεται στο σχήμα 1.5

Το κατώφλι θ_k είναι εξωτερική παράμετρος του νεύρωνα K . Έτσι μπορούμε να γράψουμε

τις (1.1), (1.2) ως εξής:

$$v_k = \sum_{j=0}^p w_{kj} \quad \text{και} \quad y_k = \phi(v_k)$$

Στην (1.4) προσθέσαμε την synapse

$$x_0 = -1 \quad \text{με} \quad \text{βάρος} \quad w_{k0} = \theta_k$$

Έτσι μπορούμε να πάρουμε του νεύρωνα K όπως φαίνεται στα σχήματα 1.6a, 1.6b που είναι ισοδύναμα με το σχήμα 1.4.

1.4 ΤΥΠΟΙ ΣΥΝΑΡΤΗΣΕΩΝ ΕΝΕΡΓΟΠΟΙΗΣΗΣ

Η συνάρτηση ενεργοποίησης $\phi(\)$, ορίζει την έξοδο ενός νεύρωνα συναρτήσει του επιπέδου ενεργοποίησης της εισόδου. Έχουμε 3 βασικούς τύπους συναρτήσεων ενεργοποίησης.

1. Συνάρτηση Κατωφλιού. Περιγράφεται στο σχήμα 1.7a και έχουμε:

$$\phi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases}$$

Επομένως η έξοδος έχει τη μορφή

$$y_k = \begin{cases} 1, & v_k \geq 0 \\ 0, & v_k < 0 \end{cases}$$

όπου v_k είναι το εσωτερικό επίπεδο ενεργοποίησης του νεύρωνα.

$$v_k = \sum_{j=0}^p w_{kj}$$

2. Τμηματικά Γραμμικές Συναρτήσεις. Περιγράφεται στο σχήμα 1.7b και έχουμε

$$\phi(v) = \begin{cases} 1, v \geq 1/2 \\ v, 1/2 > v > -1/2 \\ 0, v \leq -1/2 \end{cases}$$

Οι ακόλουθες 2 καταστάσεις μπορούν να θεωρηθούν σαν ειδικές μορφές τμηματικά γραμμικής συνάρτησης:

(α) Προσεγγίζεται ένας γραμμικός συνδυαστής, αν η γραμμική περιοχή διατηρείται σταθερή

(β) χωρίς να φτάνει σε κορεσμό. Πέφτει σε συνάρτηση κατωφλιού, αν ο ενισχυτικός παράγοντας της γραμμικής περιοχής γίνει υπερβολικά μεγάλος.

3. Σιγμοειδής. Είναι η πιο συνηθισμένη μορφή. Ένα παράδειγμα είναι η *λογιστική συνάρτηση*

$$\phi(v) = \frac{1}{1 + \exp(-av)}$$

Μεταβάλλοντας το a παίρνουμε διαφορετικές συναρτήσεις, όπως φαίνεται στο σχήμα 1.7c.

Οι προηγούμενες συναρτήσεις ενεργοποίησης κυμαίνονται από 0 ως +1. Αν θέλουμε συνάρτηση που να κυμαίνεται από -1 ως +1, επαναπροσδιορίζουμε την συνάρτηση κατωφλιού ως

$$\phi(v) = \begin{cases} 1, v > 0 \\ 0, v = 0 \\ -1, v < 0 \end{cases}$$

που ονομάζεται *signum* συνάρτηση.

1.5 ΘΕΩΡΩΝΤΑΣ ΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΩΣ ΚΑΤΕΥΘΥΝΟΜΕΝΟΥΣ ΓΡΑΦΟΥΣ

Απλοποιούμε εδώ την εμφάνιση των μοντέλων νευρώνων, χρησιμοποιώντας την ιδέα των γράφων ροής σημάτων, χωρίς να θυσιάζουμε τίποτα από τις λειτουργικές λεπτομέρειες του μοντέλου. Ένας γράφος ροής σημάτων, είναι ένα δίκτυο με κατευθυνόμενα κλαδιά, που συνδέονται σε συγκεκριμένα σημεία τους "κόμβους". Ένας τυπικός κόμβος j σχετίζεται με ένα σήμα κόμβου x_j

Ακολουθούνται 3 βασικοί κανόνες:

ΚΑΝΟΝΑΣ 1. Ένα σήμα ρέει κατά μήκος της σύνδεσης μόνο στην κατεύθυνση που ορίζεται στη σειρά.

Υπάρχουν 2 τύποι σύνδεσης:

1. *Συνδέσεις synapctic*, όπου το σήμα x_j πολλαπλασιάζεται με το βάρος w_{kj} για να παράγει το σήμα y_k , όπως φαίνεται στο σχήμα 1.8a.

2. *Συνδέσεις ενεργοποίησης*, που περιγράφει μια $\phi(\)$, που δεν είναι γραμμική και φαίνεται στο σχήμα 1.8b.

ΚΑΝΟΝΑΣ 2. Το σήμα κόμβου, ισούται με το αλγεβρικό άθροισμα όλων των σημάτων που φτάνουν στον κόμβο, σχήμα 1.8c.

ΚΑΝΟΝΑΣ 3. Το σήμα σε έναν κόμβο μεταβιβάζεται σε κάθε εξερχόμενη σύνδεση που ξεκινά από αυτόν, με την μεταβίβαση να είναι ανεξάρτητη από τις συναρτήσεις μεταφοράς των εξερχόμενων συνδέσεων, σχήμα 1.8d. Στο σχήμα 1.9 περιγράφεται το διάγραμμα του σχήματος 1.6a.

Μπορούμε τώρα να αναφέρουμε τους παρακάτω μαθηματικούς ορισμούς ενός νευρωνικού δικτύου :

Ένα νευρωνικό δίκτυο είναι ένας κατευθυνόμενος γράφος, που αποτελείται από κόμβους με synaptic διασυνδέσεις και συνδέσεις ενεργοποίησης και χαρακτηρίζονται από τις ακόλουθες ιδιότητες :

Κάθε νεύρωνας, παριστάνεται από ένα σύνολο γραμμικών synaptic συνδέσεων, ένα εξωτερικά εφαρμοζόμενο κατώφλι και μια μη-γραμμική σύνδεση ενεργοποίησης. Το κατώφλι παριστάνεται από synaptic συνδέσεις με σήμα εισόδου σταθερής τιμής -1.

Οι synaptic συνδέσεις ενός νεύρωνα ζυγίζουν τα αντίστοιχα σήματα εισόδου

Το άθροισμα των βαρών των σημάτων εισόδου καθορίζει το συνολικό εσωτερικό επίπεδο ενεργοποίησης του νεύρωνα που ζητείται.

Η σύνδεση ενεργοποίησης συνθλίβει το εσωτερικό επίπεδο ενεργοποίησης, για την παραγωγή της εξόδου που παριστάνει την κατάσταση του νεύρωνα.

Υπάρχουν όμως και μερικώς ολοκληρωμένοι κατευθυνόμενοι γράφοι που είναι γνωστοί

σαν " αρχιτεκτονικοί γράφοι". Ένας τέτοιος γράφος παριστάνεται στο σχήμα 1.10.

1.6 ΑΝΑΔΡΑΣΗ

Η Ανάδραση υπάρχει σε ένα δυναμικό σύστημα κάθε φορά που η έξοδος ενός στοιχείου επηρεάζει κατά κάποιο τρόπο την είσοδο σ' αυτό το στοιχείο. Από το σχήμα 1.11 εξάγουμε τις σχέσεις εισόδου-εξόδου :

$$y_k(n) = A[x'_j(n)] \quad \text{και} \quad x'_j = x_j(n) + B[y_k(n)]$$

Έτσι παίρνουμε
$$y_k = \frac{A}{1-AB}[x_j(n)]$$

όπου ο παράγοντας $\frac{A}{1-AB}$ καλείται closed-loop τελεστής ενώ ο AB open-loop τελεστής.

1.7 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΔΙΚΤΥΩΝ

Ο τρόπος με τον οποίο οι νεύρωνες ενός νευρωνικού δικτύου δομούνται, είναι στενά συνδεδεμένα με τον αλγόριθμο εκμάθησης που χρησιμοποιείται για το train του δικτύου. Μπορούμε να διακρίνουμε 4 διαφορετικές κλάσεις αρχιτεκτονικών δομών :

Ενός-επιπέδου Επανατροφοδότησης Δίκτυα. Σε ένα τέτοιο δίκτυο, οι νεύρωνες είναι οργανωμένοι σε μορφή επιπέδων. Οι νεύρωνες του πηγαίου επιπέδου δείχνουν στους νεύρωνες του επόμενου επιπέδου αλλά όχι αντίστροφα, σχήμα 1.14.

Πολλαπλών-επιπέδων Επανατροφοδότησης Δίκτυα. Εδώ έχουμε περισσότερα του ενός κρυφά επίπεδα, των οποίων οι κόμβοι υπολογισμού ονομάζονται "κρυφοί νεύρωνες". Τυπικά, οι νεύρωνες σε κάθε επίπεδο έχουν σαν εισόδους τα σήματα εξόδου του προηγούμενου μόνο επιπέδου.

Στο σχήμα 1.15 έχουμε ένα ολικά-συνδεδεμένο νευρωνικό δίκτυο, με την έννοια ότι κάθε κόμβος να συνδέεται με όλους τους κόμβους του αμέσως επόμενου επιπέδου. Αντίθετα στο σχήμα 1.16 έχουμε ένα μερικώς-συνδεδεμένο νευρωνικό δίκτυο.

Επαναλαμβανόμενα Δίκτυα. Η διαφορά με τα Δίκτυα Επανατροφοδότησης είναι ότι έχει ένα τουλάχιστον βρόχο ανάδρασης. Στο σχήμα 1.18 φαίνεται ένα Επαναλαμβανόμενο δίκτυο με κρυφούς νεύρωνες.

Δικτυωτές Δομές. Ένα κυκλίδωμα, αποτελείται από έναν πίνακα μιας-δύο ή μεγαλύτερης διάστασης από νεύρωνες, με ένα αντίστοιχο σύνολο από πηγαίους κόμβους, που παρέχουν τα σήματα εισόδου στον πίνακα.

1.8 ΑΝΑΠΑΡΑΣΤΑΣΗ ΓΝΩΣΗΣ

Στην παράγραφο 1.1. χρησιμοποίησαμε τον όρο "γνώση" στον ορισμό ενός νευρωνικού δικτύου χωρίς μία ιδιαίτερη περιγραφή του τι εννοούμε με αυτό. Τώρα ενδιαφερόμαστε γι' αυτό το θέμα δίνοντας τον ακόλουθο γενικό ορισμό:

Με τον όρο γνώση αναφερόμαστε σε αποθηκευμένη πληροφορία ή σε μοντέλα χρησιμοποιούμενα από ένα άτομο ή μηχανή για να μεταφράσουν, προβλέψουν και κατά προσέγγιση να αντιδράσουν στον εξωτερικό κόσμο.

Δίκτυα αναπαριστούν μια ιδιαίτερη ομάδα έξυπνων μηχανών. Τυπικά παρ' όλα αυτά οι πιθανοί τύποι αναπαράστασης από τις εισόδους προς τις εσωτερικές παραμέτρους του δικτύου ποικίλουν πάρα πολύ, πράγμα το οποίο τείνει να κάνει την ανάπτυξη μιας ικανοποιητικής λύσης μέσω ενός νευρωνικού δικτύου μια πραγματική σχεδιαστική πρόκληση.

Η μέγιστη εργασία για ένα νευρωνικό δίκτυο είναι να μάθει ένα μοντέλο του κόσμου (περιβάλλον) στο οποίο είναι εγκατεστημένο και να συντηρήσει το μοντέλο ικανοποιητικά σύμφωνα με τον πραγματικό κόσμο, ώστε να επιτύχει τους στόχους μιας εφαρμογής που ενδιαφερόμαστε. Η γνώση του κόσμου αποτελείται από δύο είδη πληροφορίας:

1. *Η γνωστή κατάσταση του κόσμου*, αναπαριστόμενη από παράγοντες σχετικά με το τι είναι και ήταν γνωστό. Αυτός ο τύπος γνώσης αναφέρεται και σαν από πριν (prior) πληροφορία.

2. *Παρατηρήσεις* (μετρήσεις) του κόσμου, που έχουν αποκτηθεί από αισθητήρες σχεδιασμένους να διερευνούν το περιβάλλον στο οποίο το νευρωνικό δίκτυο υποτίθεται ότι λειτουργεί. Συνήθως αυτές οι παρατηρήσεις είναι κληρονομικά θορυβώδεις καθώς υπόκεινται σε λάθη εξαιτίας των θορύβων του αισθητήρα και των ατελειών του συστήματος. Σε οποιαδήποτε περίπτωση, οι παρατηρήσεις που αποκτώνται με αυτό τον τρόπο δίνουν το κυρίως μέρος των πληροφοριών, από τις

οποίες παίρνονται τα παραδείγματα τα οποία χρησιμοποιούνται για την εκπαίδευση του νευρωνικού δικτύου.

Κάθε παράδειγμα αποτελείται από ένα ζεύγος εισόδου/εξόδου : ένα σήμα εισόδου και η αντίστοιχη επιθυμητή αντίδραση για το νευρωνικό δίκτυο. Γι'αυτό ένα σύνολο παραδειγμάτων αναπαριστά γνώση σχετικά με το περιβάλλον ενδιαφέροντος. Θεωρείστε για παράδειγμα το πρόβλημα της *χειρόγραφης αναγνώρισης ψηφίων*, στο οποίο η είσοδος αποτελείται από μία εικόνα από άσπρα και μαύρα pixels και κάθε εικόνα αναπαριστά ένα από τα δέκα ψηφία τα οποία διαχωρίζονται από το φόντο. Σ'αυτό το παράδειγμα η επιθυμητή αντίδραση ορίζεται από την "*ταυτότητα*" ενός συγκεκριμένου ψηφίου του οποίου η εικόνα αναπαρίσταται στο νευρωνικό δίκτυο σαν ένα σήμα εισόδου. Τυπικά το σύνολο των παραδειγμάτων που χρησιμοποιείται για την εκπαίδευση του δικτύου αποτελείται από μια μεγάλη ποικιλία ψηφίων που είναι αντιπροσωπευτικά της πραγματικής εικόνας του δικτύου. Δοσμένου ενός τέτοιου συνόλου παραδειγμάτων η σχεδίαση ενός νευρωνικού δικτύου μπορεί να γίνει ως εξής:

Πρώτα μια *κατάλληλη αρχιτεκτονική* επιλέγεται για το νευρωνικό δίκτυο, με μία είσοδο αποτελούμενη από κόμβους εισόδου ισάριθμους προς τα pixel μιας εικόνας εισόδου και μια έξοδο αποτελούμενη από δέκα νευρώνες (έναν για κάθε ψηφίο). Ένα υποσύνολο παραδειγμάτων χρησιμοποιείται τότε για την εκπαίδευση του δικτύου μέσω ενός κατάλληλου αλγορίθμου. Αυτή η φάση λέγεται μάθηση.

Δεύτερο η απόδοση της αναγνώρισης του εκπαιδευμένου δικτύου τεστάρεται με δεδομένα που δεν έχουν εισαχθεί πριν. Συγκεκριμένα μια εικόνα εισόδου παρουσιάζεται στο δίκτυο αλλά αυτή τη φορά δεν του λέγεται η ταυτότητα του ψηφίου στο οποίο ανήκει αυτή η συγκεκριμένη εικόνα. Η απόδοση του δικτύου τότε παίρνεται συγκρίνοντας την αναγνώριση του ψηφίου που αναφέρεται από το δίκτυο με την πραγματική ταυτότητα του ψηφίου.

Αυτή η δεύτερη φάση της λειτουργίας του δικτύου ονομάζεται *γενίκευση*, ένας όρος που είναι δανεισμένος από την ψυχολογία. Σ'αυτό βρίσκεται μια θεμελιώδης διαφορά ανάμεσα στον σχεδιασμό ενός νευρωνικού δικτύου και στο μέρος της κλασσικής επεξεργασίας πληροφορίας (*pattern classifier*). Στην τελευταία περίπτωση συνήθως προχωράμε σχηματίζοντας πρώτα ένα μαθηματικό μοντέλο περιβαλλοντικών παρατηρήσεων επαληθεύοντας το μοντέλο με πραγματικά δεδομένα και κατόπιν χτίζοντας τον σχεδιασμό από τη βάση του μοντέλου. Σε αντίθεση ο σχεδιασμός ενός νευρωνικού δικτύου βασίζεται απ'ευθείας σε πραγματικά δεδομένα με το σύνολο των δεδομένων να κάνει απ' ευθείας όλη τη δουλειά. Γι' αυτό το νευρωνικό δίκτυο όχι μόνο δίνει ένα τέλειο μοντέλο του περιβάλλοντος στο οποίο εγκαθίσταται, αλλά παρουσιάζει μια εργασία επεξεργασίας πληροφοριών μεγάλου ενδιαφέροντος.

Τα παραδείγματα που χρησιμοποιούνται για την εκπαίδευση ενός νευρωνικού δικτύου μπορεί να αποτελούνται και από *θετικά* και από *αρνητικά* παραδείγματα. Για παράδειγμα στο πρόβλημα της παθητικής ανίχνευσης με σόναρ για θετικά παραδείγματα δίνονται σαν είσοδο δεδομένα εκπαίδευσης που περιέχουν τον στόχο ενδιαφέροντος (π.χ. υποβρύχιο). Τώρα σ'ένα παθητικό περιβάλλον σόναρ η πιθανή παρουσία υποβρύχιας ζωής στα δεδομένα τεσταρίσματος προκαλεί περιστασιακούς λάθος συναγερμούς. Για να ανακουφιστεί αυτό το πρόβλημα αρνητικά παραδείγματα (π.χ. ηχώ από υποβρύχια ζωή) περιλαμβάνονται στα δεδομένα εκπαίδευσης ώστε να μάθει το δίκτυο να μην μπερδεύει την υποβρύχια ζωή με το στόχο.

Σ' ένα νευρωνικό δίκτυο μιας συγκεκριμένης αρχιτεκτονικής η αναπαράσταση γνώσης του περιβάλλοντος ορίζεται από τις τιμές που παίρνουν οι ελεύθεροι

παράμετροι (π.χ. συναπτικά βάρη και αδιέξοδα) του δικτύου. Ο τύπος αυτής της αναπαράστασης γνώσης αποτελεί τους διαφορετικούς σχεδιασμούς ενός νευρωνικού δικτύου και γι' αυτό κρατάει το κλειδί της απόδοσης του.

Το θέμα της αναπαράστασης γνώσης μέσα σ' ένα τεχνητό νευρωνικό δίκτυο είναι παρ' όλα αυτά πολύ περιπλεγμένο. Το θέμα γίνεται ακόμα πιο πολύπλοκο όταν έχουμε πολλαπλές πηγές πληροφορίας να ενεργούν πάνω στο δίκτυο και αυτές οι πηγές αλληλεπιδρούν μεταξύ τους. Η παρούσα κατανόηση αυτού του σημαντικού θέματος είναι πραγματικά η πιο αδύναμη σύνδεση πάνω σ' αυτά που ξέρουμε σχετικά με τεχνητά νευρωνικά δίκτυα. Παρόλα αυτά υπάρχουν 4 κανόνες για αναπαράσταση γνώσης που είναι κοινής λογικής. Οι κανόνες αυτοί περιγράφονται παρακάτω:

ΚΑΝΟΝΑΣ 1: Παρόμοιες εισόδοι από παρόμοιες ομάδες πρέπει συνήθως να παράγουν παρόμοιες αναπαραστάσεις μέσα στο δίκτυο και πρέπει γι' αυτό να ταξινομούνται σαν να ανήκουν στην ίδια κατηγορία.

Υπάρχει μια πληθώρα μέτρων για να αποφασιστεί η "ομοιότητα" ανάμεσα σε εισόδους. Ένα συνηθισμένο μέτρο ομοιότητας βασίζεται στην έννοια της Ευκλείδειας απόστασης. Για να γίνουμε πιο συγκεκριμένοι έστω x_i ένα N-διάστατο πραγματικό διάνυσμα

$$x_i = [x_{i1}, x_{i2}, \dots, x_{iN}]^T \quad (1.23)$$

που όλα τα στοιχεία είναι πραγματικά, το T δηλώνει μετατόπιση πίνακα. Το διάνυσμα x_j ορίζει ένα σημείο σ' ένα N-διάστατο χώρο που καλείται Ευκλείδειος χώρος και συμβολίζεται με R^N . Η απόσταση μεταξύ x_i και x_j ορίζεται σαν

$$d_{ij} = \|x_i - x_j\| = \left[\sum_{n=1}^N (x_{in} - x_{jn})^2 \right]^{1/2} \quad (1.24)$$

όπου x_{in} και x_{jn} είναι τα N-οστά στοιχεία του x_i και x_j . Αντίστοιχα η ομοιότητα ανάμεσα στις εισόδους που αναπαρίστανται από τα διανύσματα x_i και x_j ορίζεται σαν η παλινδρόμηση της Ευκλείδειας απόστασης d_{ij} . Όσο πιο κοντά είναι τα ξεχωριστά στοιχεία των διανυσμάτων εισόδου x_i και x_j τόσο μικρότερη θα είναι η d_{ij} και τόσο μεγαλύτερη θα είναι η ομοιότητα των x_i και x_j . Ο κανόνας 1 δηλώνει ότι αν τα διανύσματα x_i και x_j είναι παρόμοια τότε θα πρέπει να τοποθετηθούν στην ίδια κατηγορία.

Ένα άλλο μέτρο της ομοιότητας βασίζεται στην ιδέα του *dot product* ή *εσωτερικού γινομένου* που το δανειζόμαστε από την γραμμική άλγεβρα. Δοθέντων δύο διανυσμάτων x_i και x_j ίδιας διάστασης το εσωτερικό του γινόμενο ορίζεται σαν

$$x_i^T x_j = \sum_{n=1}^N x_{in} x_{jn} \quad (1.25)$$

Τα μέτρα ομοιότητας πραγματικά σχετίζονται μεταξύ τους όπως φαίνεται στο σχήμα 1.20. Η Ευκλείδεια απόσταση $\|x_i - x_j\|$ ανάμεσα στα δύο διανύσματα x_i και x_j σκισσάρεται σαν το μήκος της γραμμής που ενώνει τα άκρα των δύο αυτών διανυσμάτων, και το εσωτερικό του γινόμενο σκισσάρεται σαν την προβολή του x_i πάνω στο διάνυσμα x_j . Το σχήμα 1.20. δείχνει καθαρά ότι όσο μικρότερη είναι η Ευκλείδεια απόσταση, και γι' αυτό είναι ποιο όμοια τα x_i και x_j , τόσο μεγαλύτερο θα είναι το εσωτερικό τους γινόμενο. Με όρους επεξεργασίας σημάτων το εσωτερικό γινόμενο μπορεί να ειπωθεί σαν μια *συνάρτηση cross-correlation*. Αναγνωρίζοντας ότι το εσωτερικό γινόμενο είναι βαθμωτό μπορούμε να δηλώσουμε ότι όσο πιο θετικό είναι τόσο πιο όμοια είναι τα διανύσματα x_i και x_j . Η συνάρτηση *cross-correlation* είναι ιδανική για εύρεση ηχούς σε ραντάρ και σόναρ συστήματα. Συγκεκριμένα περνώντας

την ηχώ με συνάρτηση cross-correlate απ'το στόχο με ένα αντίγραφο του μεταδιδόμενου σήματος και βρίσκοντας την κορυφαία τιμή της συνάρτησης αποτελέσματος, είναι εύκολο να υπολογιστεί ο χρόνος φτασίματος της ηχούς. Αυτή είναι βασική μέθοδος υπολογισμού της απόστασης του στόχου.

ΚΑΝΟΝΑΣ 2 : Αντικείμενα που κατηγοριοποιούνται σε διαφορετικές ομάδες θα πρέπει να έχουν διαφορετικές αναπαραστάσεις στο δίκτυο. Ο δεύτερος κανόνας είναι αντίθετος του πρώτου.

ΚΑΝΟΝΑΣ 3 : Αν ένα συγκεκριμένο χαρακτηριστικό είναι σημαντικό τότε θα πρέπει να υπάρχει ένας μεγάλος αριθμός από νευρώνες που συμμετάσχουν στην αναπαράσταση αυτού του αντικειμένου στο δίκτυο.

Θεωρείστε για παράδειγμα την εφαρμογή ραντάρ για την ανίχνευση ενός στόχου με την παρουσία θορύβου. Σύμφωνα με το κριτήριο *Neyman-Pearson* η πιθανότητα ανίχνευσης μεγιστοποιείται αν το όριο της πιθανότητας λάθος συναγερμού δεν υπερβαίνει μια συγκεκριμένη τιμή. Σε μια τέτοια εφαρμογή η πραγματική παρουσία ενός στόχου στο λαμβανόμενο σήμα αναπαριστά ένα σημαντικό χαρακτηριστικό της εισόδου. *Ο κανόνας 3 σαν αποτέλεσμα δηλώνει ότι θα πρέπει να υπάρχει ένας πολύ μεγάλος αριθμός νευρώνων ώστε να αποφασιστεί ότι ο στόχος είναι παρών όταν πράγματι είναι.* Κατά τον ίδιο τρόπο θα πρέπει να υπάρχει ένας πολύ μεγάλος αριθμός νευρώνων ώστε να αποφασιστεί ότι υπάρχει μόνο θόρυβος. Και στις δύο περιπτώσεις ο μεγάλος αριθμός νευρώνων βεβαιώνει το μεγάλο βαθμό ακρίβειας της απόφασης και της απόκλισης εξαιτίας λανθασμένων νευρώνων.

ΚΑΝΟΝΑΣ 4 : Η από πριν (prior) πληροφορίες και μη μεταβλητότητες θα πρέπει να χτιστούν στον σχεδιασμό ενός νευρωνικού δικτύου απλοποιώντας έτσι τον σχεδιασμό αυτό του δικτύου ,το οποίο δεν έχει να τις μάθει.

Ο κανόνας 4 είναι ιδιαίτερα σημαντικός διότι η κληρονομικότητα των αποτελεσμάτων τους δίνει ένα νευρωνικό δίκτυο με μια εξειδικευμένη δομή.Αυτό είναι ιδιαίτερα επιθυμητό για διάφορους λόγους:

- 1) Βιολογικά , οπτικά και ελεγκτικά δίκτυα είναι γνωστό ότι είναι πολύ εξειδικευμένα.
- 2) Ένα νευρωνικό δίκτυο με εξειδικευμένη δομή συνήθως έχει πολύ μικρότερο αριθμό ελεύθερων παραμέτρων διαθέσιμων για κανονικοποίηση απ' ότι ένα πλήρως συνδεδεμένο δίκτυο. Συνεπώς ένα εξειδικευμένο δίκτυο απαιτεί λιγότερα δεδομένα εισόδου για εκπαίδευση, μαθαίνει γρηγορότερα και συχνά γενικοποιεί καλύτερα.
- 3) Ο ρυθμός της μετάδοσης πληροφοριών μέσω ενός εξειδικευμένου δικτύου επιταχύνεται.
- 4) Το κόστος χτισίματος ενός εξειδικευμένου δικτύου μειώνεται εξαιτίας του μικρότερου μεγέθους σε σύγκριση μ' αυτό του πλήρως συνδεδεμένου.

1.9 ΠΩΣ ΝΑ ΧΤΙΣΕΤΕ PRIOR ΠΛΗΡΟΦΟΡΙΕΣ ΜΕΣΑ ΣΤΟ ΣΧΕΔΙΑΣΜΟ ΕΝΟΣ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

Ένα σημαντικό είναι το πως να αναπτύξεις μια εξειδικευμένη δομή χτίζοντας prior πληροφορίες στο σχεδιασμό του δικτύου. Δυστυχώς δεν υπάρχουν καλά ορισμένοι κανόνες για να το κάνεις αυτό, παρόλο που έχουμε μερικές *ad hoc διαδικασίες* που βγάζουν χρήσιμα συμπεράσματα.

Για να γίνουμε πιο συγκεκριμένοι θεωρείστε το πρόβλημα της αναγνώρισης χειρόγραφων ψηφίων που είναι μια σχετικά εύκολη ανθρώπινη δουλειά αλλά καθόλου εύκολη για τη μηχανή. Η είσοδος αποτελείται από ασπρόμαυρα pixels που αναπαριστούν ένα από τα δέκα ψηφία τα οποία ξεχωρίζονται από τον φόντο. Σ' αυτό το παράδειγμα η prior πληροφορία είναι ότι η εικόνα είναι δισδιάστατη και έχει μια ισχυρή τοπική δομή. Γι' αυτό το δίκτυο εξειδικεύεται περιορίζοντας τις συναπτικές συνδέσεις στις πρώτες υποδοχές του δικτύου ώστε να είναι τοπικές, δηλαδή το δίκτυο επιλέγεται να είναι τοπικά συνδεδεμένο.

Επιπλέον εξειδίκευση μπορεί να γίνει στο σχεδιασμό του δικτύου, εξετάζοντας τη χρήση ενός *ανιχνευτή χαρακτηριστικών* ο οποίος οδηγεί στη μείωση δεδομένων εισόδου επιλέγοντας ορισμένα χαρακτηριστικά που ξεχωρίζουν την εικόνα του ενός ψηφίου απ' το άλλο. Συγκεκριμένα αν ένας ανιχνευτής είναι χρήσιμος σε ένα μέρος της εικόνας τότε είναι πιθανόν να είναι χρήσιμος και σε άλλο. Ο λόγος που το λέμε αυτό είναι ότι τα προεξέχοντα χαρακτηριστικά ενός χαρακτήρα μπορεί να απέχουν ελαφρά από τη θέση τους. Για να λυθεί αυτό το πρόβλημα η εικόνα εισόδου ανιχνεύεται από έναν απλό νευρώνα που έχει ένα τοπικό επιδεκτικό πεδίο και τα συναπτικά βάρη του νευρώνα αποθηκεύονται σε αντίστοιχες θέσεις σε μια σύνδεση που λέγεται χάρτης χαρακτηριστικών. Η λειτουργία φαίνεται στο σχήμα 1.21. Εστω $\{ w_{ji} | i = 0, 1, \dots, p-1 \}$ δηλώνει το σύνολο των συναπτικών βαρών του νευρώνα j . Η περιέλιξη του πυρήνα αναπαριστάτε από αυτό το σύνολο των συναπτικών βαρών και ένα pixel εισόδου που δηλώνεται σαν $\{x(n)\}$ ορίζεται από το άθροισμα

$$y_j(n) = \sum_{i=0}^{p-1} w_{ji} x(n-i) \quad (1.26)$$

Όπου το n δηλώνει το n -ιοστο δείγμα ενός pixel εισόδου, ώστε το δίκτυο καλείται μερικές φορές *δίκτυο περιέλιξης*. Γι' αυτό η συνολική λειτουργία που φαίνεται στο σχήμα 1.21 είναι ισοδύναμη προς την περιέλιξη ενός μικρού μεγέθους πυρήνα του νευρώνα και της εικόνας εισόδου που τότε ακολουθείται από έναν μικρό περιορισμό που γίνεται από την συνάρτηση ενεργοποίησης του νευρώνα. Η συνολική λειτουργία γίνεται παράλληλα σ' ένα επίπεδο από νευρώνες των οποίων τα διανύσματα βαρών είναι ίσα. Με άλλα λόγια οι νευρώνες ενός χάρτη χαρακτηριστικών περιορίζονται ώστε να παρουσιάσουν την ίδια μαθηματική λειτουργία σε διαφορετικά μέρη της εικόνας. Μια τέτοια τεχνική καλείται διαμοίραση βαρών. Η διαμοίραση βαρών έχει επίσης μια επικερδή παρενέργεια: Ο αριθμός των ελεύθερων παραμέτρων του δικτύου μειώνεται σημαντικά καθώς ένας μεγάλος αριθμός από νευρώνες στο δίκτυο περιορίζεται ώστε να μοιράζεται το ίδιο σύνολο συναπτικών βαρών.

Περίληπτικά η prior πληροφορία μπορεί να χτιστεί στον σχεδιασμό ενός νευρωνικού δικτύου μέσω ενός συνδυασμού δύο τεχνικών: (1) περιορίζοντας την αρχιτεκτονική του δικτύου μέσω της χρήσης τοπικών συνδέσεων και (2) περιορίζοντας την επιλογή των συναπτικών βαρών χρησιμοποιώντας διαμοίραση βαρών. Φυσικά ο τρόπος με τον οποίο αυτές οι δύο τεχνικές γίνονται στην πράξη επηρεάζεται σημαντικά απ' την εφαρμογή. Πιο γενικά η ανάπτυξη καλά ορισμένων διαδικασιών για την χρήση prior πληροφορίας είναι ένα ανοιχτό πρόβλημα. Η prior πληροφορία κυριαρχεί στο ένα μέρος του κανόνα 4, το υπόλοιπο μέρος αναφέρεται στην μη μεταβλητότητα.

1.9.1 ΠΩΣ ΧΤΙΖΕΤΑΙ ΜΗ ΜΕΤΑΒΛΗΤΟΤΗΤΑ ΣΤΟ ΣΧΕΔΙΑΣΜΟ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Όταν ένα αντικείμενο περιστρέφεται, η εικόνα του όπως φαίνεται από έναν παρατηρητή συνήθως αλλάζει με αντίστοιχο τρόπο. Σε ένα ραντάρ που δίνει το πλάτος καθώς και πληροφορίες σχετικά με τη φάση για το περιβάλλον η ηχώ από έναν κινούμενο στόχο ανεβαίνει σε συχνότητα εξαιτίας του αποτελέσματος Doppler, που προκύπτει εξαιτίας της κυκλικής κίνησης του στόχου σε σχέση με το ραντάρ. Η έκφραση ενός ατόμου μπορεί να γίνει με χαμηλή ή υψηλή φωνή και παρόλα αυτά με γρήγορο ή αργό τρόπο. Για να χτιστεί ένα σύστημα αναγνώρισης αντικειμένου, αναγνώρισης στόχου και ένα σύστημα αναγνώρισης ομιλίας αντίστοιχα το σύστημα πρέπει να είναι ικανό να ανταπεξέλθει σε ένα μήκος παραμορφώσεων του παρατηρούμενου σήματος. Αντίστοιχα μια πρωτεύουσα απαίτηση ενός σχεδίου αναγνώρισης είναι να σχεδιαστεί ένας ταξινομητής που είναι αμετάβλητος σε τέτοιες παραμορφώσεις. Με άλλα λόγια ένα όριο τάξης που αναπαρίσταται από μια έξοδο του ταξινομητή δεν θα πρέπει να επηρεάζεται από παραμορφώσεις του παρατηρούμενου σήματος εισόδου.

Υπάρχουν τρεις τεχνικές για να γίνει αυτό:

1) *Μη μεταβλητότητα μέσω δομής*: Η μη μεταβλητότητα μπορεί να τοποθετηθεί σε ένα νευρωνικό δίκτυο δομώντας το σχηματισμό του κατάλληλα.

Συγκεκριμένα συναπτικές συνδέσεις ανάμεσα στους νευρώνες του δικτύου δημιουργούνται ώστε παραμορφωμένες εκδόσεις της ίδιας εισόδου να παράγουν την ίδια έξοδο.

Θεωρείστε για παράδειγμα την κατηγοριοποίηση μιας εικόνας εισόδου από ένα νευρωνικό δίκτυο, ώστε να απαιτείται να είναι ανεξάρτητη από τις εντός επιπέδου περιστροφές σε σχέση με το κέντρο της εικόνας. Μπορούμε να παράγουμε περιστροφική αμεταβλητότητα στη δομή του δικτύου όπως ακολουθεί.

Έστω w_{ij} είναι το συναπτικό βάρος του νευρώνα j που συνδέεται με το pixel i στην εικόνα εισόδου. Αν η συνθήκη $w_{ji} = w_{jk}$ επιβάλλεται για όλα τα pixels i και k που βρίσκονται σε ίσες αποστάσεις από το κέντρο της εικόνας, τότε το νευρωνικό δίκτυο είναι αμετάβλητο σε εντός επιπέδου περιστροφές. Παρόλα αυτά ώστε να διατηρηθεί η περιστροφική αμεταβλητότητα το συναπτικού βάρους w_{ji} θα πρέπει να ανατυπωθεί για κάθε pixel της εικόνας εισόδου στην ίδια ακτινική απόσταση από το πρωτότυπο. Αυτό δείχνει μια μεταβλητότητα της δομής: ο αριθμός των συναπτικών συνδέσεων στο δίκτυο γίνεται αρκετά μεγάλος ακόμα και για εικόνες μετρίου μεγέθους.

2) *Αμεταβλητότητα μέσω εκπαίδευσης*: Ένα νευρωνικό δίκτυο έχει την φυσική ικανότητα ταξινόμησης προτύπων. Αυτή η ικανότητα μπορεί να επιβληθεί απευθείας ώστε να αποκτηθεί μη μεταβλητότητα παραμορφώσεων όπως ακολουθεί. Το δίκτυο εκπαιδεύεται δίνοντάς του έναν αριθμό από διαφορετικά παραδείγματα του ίδιου αντικειμένου με τα παραδείγματα να έχουν εκλεγεί ώστε να αντιστοιχούν σε διαφορετικές παραμορφώσεις του αντικειμένου. Αρκεί ο αριθμός των παραδειγμάτων να είναι αρκετά μεγάλος και αν το δίκτυο εκπαιδεύεται ώστε να μάθει να διαχωρίζει τις διαφορετικές οπτικές γωνίες του αντικειμένου τότε μπορούμε να περιμένουμε από το δίκτυο να γενικεύει παραμορφώσεις άλλες από αυτές που του δείξαμε.

Παρόλα αυτά από μηχανικής πλευράς η μεταβλητότητα μέσω εκπαίδευσης έχει δύο μειονεκτήματα:

a) Όταν ένα νευρωνικό δίκτυο έχει εκπαιδευτεί για να αναγνωρίζει ένα αντικείμενο με αμετάβλητο τρόπο ως προς τους γνωστούς μετασχηματισμούς δεν

είναι εμφανές ότι αυτή η εκπαίδευση θα κάνει ικανό το δίκτυο να αναγνωρίσει άλλα αντικείμενα διαφορετικών ομάδων αμετάβλητα.

b) Η υπολογιστική απαίτηση στο δίκτυο μπορεί να είναι πολύ δύσκολη ειδικά αν ο βαθμός του χώρου χαρακτηριστικών είναι υψηλός.

3) *Μη μεταβλητότητα χώρου χαρακτηριστικών*: Η τρίτη τεχνική φαίνεται στο σχήμα 1.22. Βασίζεται στο γεγονός ότι μπορεί να είναι δυνατόν να παρθούν χαρακτηριστικά που χαρακτηρίζουν την βασική πληροφορία των δεδομένων εισόδου και που είναι αμετάβλητα ως προς παραμορφώσεις της εισόδου. Αν τέτοια χαρακτηριστικά χρησιμοποιούνται, τότε το δίκτυο σαν ταξινομητής ανακουφίζεται από τον μπελά του να έχει να απεικονίσει το μήκος των παραμορφώσεων ενός αντικειμένου με περίπλοκα όρια απόφασης. Πραγματικά μοναδικές διαφορές που μπορούν να προκύψουν ανάμεσα σε διαφορετικές φάσεις του ίδιου αντικειμένου είναι εξαιτίας αναπόφευκτων παραγόντων όπως ο θόρυβος.

Η χρησιμοποίηση του αμετάβλητου χώρου χαρακτηριστικών προσφέρει τρία *πλεονεκτήματα*: (1) Ο αριθμός των χαρακτηριστικών στο δίκτυο μπορεί να μειωθεί σε ρεαλιστικά επίπεδα, (2) οι απαιτήσεις που σχετίζονται με την σχεδίαση του δικτύου ελαφρύνονται και (3) η αμεταβλητότητα για όλα τα αντικείμενα ως προς τους γνωστούς μετασχηματισμούς είναι αβέβαιη, παρόλα αυτά η προσέγγιση απαιτεί από πριν γνώση του προβλήματος.

Συμπερασματικά, η χρήση ενός αμετάβλητου χώρου χαρακτηριστικών προσφέρει την πιο κατάλληλη τεχνική για νευρωνικούς ταξινομητές. Για να δειχτεί η ιδέα του αμετάβλητου χώρου χαρακτηριστικών θεωρείστε το παράδειγμα ενός συστήματος ραντάρ που χρησιμοποιείται για την επιτήρηση του αέρα όπου οι στόχοι περιλαμβάνουν αεροπλάνα, συστήματα κενού, πουλιά, αντικείμενα εδάφους. Είναι ότι η ηχώ των ραντάρ απ'αυτούς τους στόχους έχει διαφορετικά χαρακτηριστικά. Επιπλέον πειραματικές έχουν δείξει ότι τέτοια συστήματα ραντάρ μπορούν να μοντελοποιηθούν αρκετά κοντά σαν μια *autoregressive* διαδικασία. Ένα AR μοντέλο είναι ένας ιδιαίτερος τύπος *regressive* μοντέλου που ορίζεται για δεδομένα πολύπλοκων τιμών

$$x(n) = \sum_{i=1}^M a_i^* x(n-i) + e(n) \quad (1.27)$$

όπου $\{a_i \mid i = 1, 2, \dots, M\}$ είναι οι συντελεστές του AR, M είναι η τάξη του μοντέλου, όπου x(n) είναι η είσοδος και e(n) είναι το λάθος που περιγράφεται σαν λευκός θόρυβος.

Βασικά το AR μοντέλο αναπαρίσταται από ένα *φίλτρο καθυστερημένης γραμμής* σχ. 1.23α για M=2. Ισοδύναμα μπορεί να αναπαρασταθεί από ένα δικτυωτό φίλτρο όπως φαίνεται στο σχήμα 1.23β, οι συντελεστές του οποίου ονομάζονται συντελεστές ανάκλασης. Υπάρχει μια ένα προς ένα σχέση ανάμεσα στους συντελεστές AR του μοντέλου του σχ. 1.23α και των συντελεστών ανάκλασης του σχ. 1.23β. Τα μοντέλα του σχήματος 1.23 υποθέτουν ότι η είσοδος x(n) είναι σύνθετη όπως στην περίπτωση των ραντάρ, στην οποία οι συντελεστές AR και οι συντελεστές ανάκλασης είναι όλοι πολύπλοκων τιμών. Για την ώρα αρκεί να πούμε ότι τα δεδομένα του ραντάρ μπορούν να περιγραφούν από ένα σύνολο από *συντελεστές autoregressive*, ή ισοδύναμα από ένα σύνολο από συντελεστές ανάκλασης. Το τελευταίο σύνολο έχει ένα υπολογιστικό πλεονέκτημα το ότι επαρκείς αλγόριθμοι υπάρχουν για τον υπολογισμό τους απ' ευθείας από τα δεδομένα εισόδου. Το πρόβλημα εξαγωγής χαρακτηριστικών παρόλα αυτά είναι πολύπλοκο εξαιτίας του γεγονότος ότι κινούμενα αντικείμενα παράγουν ποικίλες *Doppler* που εξαρτώνται από τις ακτινικές τους ταχύτητες, μετρημένες ως προς το ραντάρ, που τείνουν να

παραμορφώσουν το περιεχόμενο του φάσματος των συντελεστών ανάκλασης, καθώς το χαρακτηριστικό μειώνεται.

Για να ανταπεξέλθουμε σ' αυτό το πρόβλημα Θα πρέπει να χτίσουμε Doppler αμεταβλητότητα στον υπολογισμό των συντελεστών ανάκλασης. Η γωνία φάσης του πρώτου συντελεστή ανάκλασης καταλήγει να είναι ίση προς τη συχνότητα Doppler του σήματος του ραντάρ. Οσον αφορά την κανονικοποίηση της συχνότητας Doppler εφαρμόζεται σε όλους τους συντελεστές ώστε να φύγουν οι περιττές μετατοπίσεις Doppler. Αυτό γίνεται με τον ορισμό ενός νέου συνόλου συντελεστών ανακλάσεων $\{\rho'm\}$ σχετισμένων με τους συντελεστές ανάκλασης $\{\rho m\}$ ώστε

$$\rho_m' = \rho_m e^{-jm\theta} \quad m=1,2,\dots,M \quad (1.28)$$

όπου θ είναι η γωνία φάσης του πρώτου συντελεστή ανάκλασης και M είναι η τάξη του AR μοντέλου. Η λειτουργία του σχ. 1.28 αναφέρεται σαν heterodyning (ετερόδουνο). Ένα σύνολο από χαρακτηριστικά ραντάρ μη μεταβλητότητας Doppler έτσι αναπαρίσταται από τους κανονικοποιημένους συντελεστές ανάκλασης $\rho'1, \dots, \rho'M$ με το $\rho'1$ να είναι ο μόνος πραγματικός συντελεστής στο σύνολο. Όπως αναφέρθηκε πριν οι κύριες κατηγορίες των στόχων του ραντάρ που μας ενδιαφέρουν είναι ο καιρός, τα πουλιά, τα αεροσκάφη και το έδαφος. Οι πρώτοι τρεις στόχοι είναι κινούμενοι, ενώ ο τελευταίος δεν είναι. Οι ετερογενείς παράμετροι των ήχων του ραντάρ από το έδαφος είναι γνωστό ότι έχουν ηχούς με παρόμοια χαρακτηριστικά προς αυτό των αεροπλάνων. Είναι επίσης γνωστό ότι μια ηχώ εδάφους μπορεί να διαχωριστεί απ'αυτή ενός αεροσκάφους εξαιτίας της μικρής μετατόπισης Doppler ταξινομητής ραντάρ περιλαμβάνει έναν επεξεργαστή όπως φαίνεται στο σχήμα 1.24 που λειτουργεί στα ταξινομημένα αποτελέσματα με σκοπό την αναγνώριση της τάξης του εδάφους. Γι'αυτό ο προεπεξεργαστής αναλαμβάνει το βγάλσιμο του χαρακτηριστικού της μετατόπισης Doppler στην είσοδο του ταξινομητή ενώ ο μεταεπεξεργαστής χρησιμοποιεί τα ταξινομημένα αποτελέσματα Doppler για να διαχωρίσει τις επιστροφές αεροπλάνων και εδάφους.

Ένα περισσότερο συναρπαστικό παράδειγμα αναπαράστασης γνώσης βρίσκεται σ'ένα νευρωνικό δίκτυο βιολογικού συστήματος ανίχνευσης για τον εντοπισμό νυχτερίδων με ηχώ. Οι περισσότερες νυχτερίδες χρησιμοποιούν FM σήματα (σ'ένα FM σήμα η συχνότητα του σήματος αλλάζει με το χρόνο). Πιο συγκεκριμένα η νυχτερίδα χρησιμοποιεί το στόμα για να μεταδώσει μικρής διάρκειας FM σήματα και χρησιμοποιεί το ακουστικό σύστημα σαν παραλήπτη. Η ηχώ από στόχους αναπαριστάται στο ακουστικό σύστημα από την ενέργεια των νευρώνων, που είναι επιλεκτικοί σε διαφορετικούς συνδυασμούς ακουστικών παραμέτρων. Υπάρχουν τρεις βασικές νευρωνικές διαστάσεις της ακουστικής αναπαράστασης της νυχτερίδας:

Συχνότητα ηχούς: που κωδικοποιείται παράγοντας τον χάρτη συχνοτήτων του κοχλίου (αυτιού), διατηρείται καθόλη της ακουστικής σαν μια συνήθης συμφωνία κατά μήκος κανονισμένων νευρώνων σε διαφορετικές συχνότητες.

Πλάτος ηχούς: που κωδικοποιείται από άλλους νευρώνες με διαφορετικά δυναμική έκταση. Ορίζεται και σαν ρύθμιση πλάτους και σαν τον αριθμό εκφορτίσεων ανά κίνητρο.

Καθυστέρηση ηχούς: που κωδικοποιείται μέσω νευρωνικών πληροφοριών που παράγουν επιλεκτικής καθυστέρησης αντιδράσεις. Ορίζεται και σαν ρύθμιση έκτασης στόχου.

Τα δύο βασικά χαρακτηριστικά μιας ηχούς στόχου για σκοπούς διαμόρφωσης εικόνας είναι το φάσμα για το σχήμα του στόχου και η καθυστέρηση του στόχου. Η νυχτερίδα παίρνει σχήμα με όρους του χρόνου φτασίματος της ηχούς από διαφορετικές ανακλώμενες επιφάνειες στο στόχο. Για να συμβεί αυτό, πληροφορίες

για τη συχνότητα στο φάσμα της ηχούς μετατρέπονται σε υπολογισμούς της δομής του χρόνου του στόχου. Πειράματα των Simmons και συνεργατών του στην μεγάλη καφέ νυχτερίδα, *Ertesicus fuscus*, δείχνουν αυτή τη μετατόπιση σαν παράλληλους μετασχηματισμούς στο χρόνο και στη συχνότητα προς χρόνο, των οποίων η έξοδος δείχνει την καθυστέρηση του στόχου. Φαίνεται ότι η ομοιομορφία οφείλεται σε συγκεκριμένες ιδιότητες των μετασχηματισμών. Επιπλέον χαρακτηριστικά μη μεταβλητότητας χτίζονται στην διαδικασία σχηματισμού ειδώλου ώστε να το κάνουν ανεξάρτητο της κίνησης του στόχου και της κίνησης της νυχτερίδας

1.10 ΔΙΑΔΙΚΑΣΙΕΣ ΑΠΕΙΚΟΝΙΣΗΣ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Μία διορατική μέθοδος για να ανταπεξέλθουμε την αδυναμία της παρούσας κατανόησής μας για την αναπαράσταση γνώσης μέσα σε ένα νευρωνικό δίκτυο, είναι να καταφύγουμε στην *πειραματική χρήση της απεικόνισης της διαδικασίας μάθησης*. Με το να το κάνουμε αυτό καταλαβαίνουμε το γεγονός ότι αναπαριστώντας πληροφοριακά δεδομένα με οπτικά μέσα, είναι η ουσία της επιστημονικής απεικόνισης. Πραγματικά μια τέτοια προσέγγιση επιτρέπει στο ανθρώπινο σύστημα εγκεφάλου-ματιού να αντιληφθεί και να συμπεράνει οπτικές πληροφορίες μέσω εικονογραφημένων μέσων, προμηθεύοντας έτσι ένα πολύ ικανό εργαλείο για τη μεταφορά πληροφορίας ανάμεσα στον εξομειωτή νευρωνικού δικτύου και του χρήστη. Έτσι δίνονται αυξημένες ευκολίες για την αλληλεπιδραστική διαχείριση της εικόνας και των διαδικασιών απεικόνισης που είναι ήδη διαθέσιμες.

Παρόλα αυτά θα πρέπει να ειπωθεί, ότι η χρήση της γραφικής αναπαράστασης της διαδικασίας μάθησης του νευρωνικού δικτύου, δεν έχει προσεχθεί όσο θα έπρεπε. Μπορούμε όμως να αναφέρουμε το *διάγραμμα Hinton* που περιγράφεται από τους Rumelhart και McClelland(1986) και το *διάγραμμα δεσμού* που προτείνεται από τους Wejchert και Tesuaro(1991). Το διάγραμμα Hinton περιέχει ζωγραφισμένες στήλες τετραγώνων, με κάθε στήλη να αναπαριστά τα συναπτικά βάρη και αδιέξοδα ενός συγκεκριμένου νευρώνα στο δίκτυο. Το μέγεθος κάθε τετραγώνου αναπαριστά το μέγεθος ενός συγκεκριμένου συναπτικού βάρους, το χρώμα (μαύρο ή άσπρο) δείχνει την πολικότητα του βάρους, θετική ή αρνητική αντίστοιχα. Οι διάφορες στήλες τοποθετούνται στο διάγραμμα ώστε να διατηρούν σχέση με την αρχιτεκτονική του δικτύου. Θεωρείστε για παράδειγμα το δίκτυο του σχ. 1.25α. Το διάγραμμα Hinton για αυτό το δίκτυο φαίνεται στο σχ. 1.25β. Αρχίζοντας από το κάτω μέρος του διαγράμματος, η πρώτη στήλη αναπαριστά τα δύο βάρη και αδιέξοδα του πάνω νευρώνα. Η δεύτερη στήλη αναπαριστά τα δύο βάρη και αδιέξοδα του κάτω νευρώνα. Τελικά η τρίτη στήλη αναπαριστά τα δύο βάρη και αδιέξοδα του νευρώνα εξόδου. Η πάνω σειρά τετραγώνων αναπαριστά τα αδιέξοδα των ξεχωριστών νευρώνων. Ένας *περιορισμός του διαγράμματος Hinton* είναι ότι δεν αποκαλύπτει την τοπολογία του δικτύου σε σχέση με τα δεδομένα των συναπτικών βαρών. Συνήθως είναι ιδιαίτερα πληροφοριακό να έχουμε μια γραφική αναπαράσταση που συνδέει τις τιμές των συναπτικών βαρών με την τοπολογία του δικτύου, γιατί τότε μπορούμε να δούμε πως η εσωτερική αναπαράσταση των συναπτικών βαρών σχετίζεται με το συγκεκριμένο πρόβλημα που το δίκτυο μαθαίνει. Ένας τέτοιος σκοπός συναντιέται στο *διάγραμμα δεσμού*, σύμφωνα με το οποίο τα συναπτικά βάρη αναπαρίστανται σαν δεσμοί ανάμεσα σε κόμβους του δικτύου. Πιο συγκεκριμένα όσο πιο δυνατή είναι μια

συναπτική σύνδεση ανάμεσα σε δύο κόμβους, τόσο πιο πολύ θα υπάρχει ο κόμβος. Το σχ.1.25c δείχνει το διάγραμμα δεσμών για το δίκτυο του σχ. 1.25a. Το σχ.1.25c δείχνει καθαρά ότι το διάγραμμα δεσμών συνδυάζει τα δεδομένα των συναπτικών βαρών με την τοπολογία του δικτύου. Προσέξτε ότι ένας δεσμός εκτείνεται και από τους δύο κόμβους της εν λόγω σύνδεσης. Γι'αυτό κοιτώντας οποιονδήποτε κόμβο στο διάγραμμα δεσμών αμέσως βλέπουμε το μέγεθος και την πολικότητα των συναπτικών βαρών που μπαίνουν και βγαίνουν απ'τον κόμβο. Αυτός ο τύπος της γραφικής αναπαράστασης προσφέρει τα ακόλουθα *πλεονεκτήματα* σε σχέση με το διάγραμμα Hinton :

Μία αίσθηση της κατεύθυνσης χτίζεται στο διάγραμμα δεσμών, καθώς κάθε δεσμός στην αναπαράσταση δείχνει στον κόμβο με τον οποίο συνδέεται.

Η συλλογή των συναπτικών βαρών που αναπαριστάται από δεσμούς, σχηματίζει ένα ιδιαίτερο σχέδιο που μπορεί να διαβαστεί από τον χρήστη, πράγμα το οποίο επιτρέπει να βγουν συμπεράσματα καθολικών πληροφοριών από το διάγραμμα δεσμών. Περιληπτικά λοιπόν το διάγραμμα δεσμών, είναι ένα οπτικό εργαλείο που αποκαλύπτει πως εσωτερικές αναπαραστάσεις στο νευρωνικό δίκτυο σχετίζονται με την λειτουργία που γίνεται στο δίκτυο. Βοηθάει επίσης στην αποκάλυψη της δυναμικής της διαδικασίας μάθησης. Παρόλο που η χρήση του διαγράμματος δεσμών αναπαριστά μια αύξηση της πολυπλοκότητας του λογισμικού, αυτό έχει μικρό κόστος σε σχέση με αυτά που κερδίζονται σχετικά με τη διαδικασία μάθησης του νευρωνικού δικτύου.

1.11 ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ ΚΑΙ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

Ένα σύστημα τεχνητής νοημοσύνης(AI) 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. Ένα σύστημα τεχνητής νοημοσύνης έχει τρία συστατικά κλειδιά: αναπαράσταση, αιτιολόγηση, και μάθηση (σχ. 1.26)

1)ΑΝΑΠΑΡΑΣΤΑΣΗ: Το πιο διαδεδομένο χαρακτηριστικό της τεχνητής νοημοσύνης πιθανόν είναι η διαδεδομένη χρήση μιας γλώσσας συμβολικών δομών για την αναπαράσταση και της γενικής γνώσης σχετικά με ένα πρόβλημα που μας ενδιαφέρει αλλά και της ειδικής γνώσης σχετικά με τη λύση του προβλήματος. Τα σύμβολα συνήθως σχηματίζονται με γνωστούς όρους , πράγμα το οποίο κάνει τις συμβολικές αναπαραστάσεις της τεχνητής νοημοσύνης σχετικά εύκολες να κατανοηθούν από τον χρήστη. Πραγματικά, η σαφήνεια της τεχνητής νοημοσύνης ευνοεί την επικοινωνία ανάμεσα σε άνθρωπο και μηχανή.

Η “γνώση” όπως χρησιμοποιείται από τους ερευνητές της τεχνητής νοημοσύνης, είναι άλλος ένας όρος για τα δεδομένα. Μπορεί να είναι *δηλωτικού ή διαδικαστικού είδους*. Σε μια *δηλωτική αναπαράσταση*, η γνώση αναπαριστάται σαν μία στατική συλλογή γεγονότων, με ένα μικρό σύνολο διαδικασιών για την διαχείριση των γεγονότων. Ένα χαρακτηριστικό στοιχείο των δηλωτικών αναπαραστάσεων είναι ότι φαίνεται να έχουν ένα δικό τους νόημα στα μάτια του χρήστη, ανεξάρτητα από τη χρήση τους στο σύστημα τεχνητής νοημοσύνης. Σε μια *διαδικαστική αναπαράσταση*, από την άλλη μεριά, η γνώση ενσωματώνεται σε έναν εκτελέσιμο κώδικα ο οποίος και δίνει το νόημα της γνώσης. Και τα δύο είδη γνώσης χρειάζονται στα περισσότερα προβλήματα που μας ενδιαφέρουν.

2) ΑΙΤΙΟΛΟΓΗΣΗ : Στην πιο βασική της μορφή, η αιτιολόγηση είναι η ικανότητα λύσης των προβλημάτων. Για να έχει ένα σύστημα την ικανότητα να είναι σύστημα αιτιολόγησης, πρέπει να ικανοποιεί ορισμένες συνθήκες:

Το σύστημα πρέπει να είναι ικανό να λύνει *ευρείας κλίμακας προβλήματα και τύπους προβλημάτων*.

Το σύστημα πρέπει να είναι ικανό να *αναγνωρίζει ρητές υπονοούμενες πληροφορίες*.

Το σύστημα πρέπει να έχει ένα *μηχανισμό ελέγχου* που να αποφασίζει ποιες λειτουργίες πρέπει να εφαρμόζονται σε ένα συγκεκριμένο πρόβλημα, όταν μια λύση του προβλήματος έχει αποκτηθεί, ή όταν η επιπλέον δουλειά στο πρόβλημα πρέπει να τερματιστεί.

Το πρόβλημα της λύσης θα πρέπει να ειπωθεί σαν ένα ερευνητικό πρόβλημα. Ένας κοινός τρόπος όταν έχεις να κάνεις με “έρευνα” είναι η *χρήση κανόνων, δεδομένων και ελέγχου*. Οι κανόνες ενεργούν πάνω στα δεδομένα και ο έλεγχος ενεργεί πάνω στους κανόνες. Θεωρείστε για παράδειγμα το πρόβλημα του πλανόδιου πωλητή, όπου η απαίτηση είναι να βρεθεί η μικρότερη διαδρομή από την μια πόλη στην άλλη, με την προϋπόθεση να πάει στις πόλεις μόνο μια φορά. Σε αυτό το πρόβλημα τα δεδομένα παίρνονται από το σύνολο των πιθανών διαδρομών και τα κόστη τους σε ένα ζυγισμένο γράφο, οι κανόνες ορίζουν τους τρόπους μετακίνησης από πόλη σε πόλη και ο έλεγχος αποφασίζει ποιοι κανόνες και πότε θα εφαρμοστούν.

Σε πολλές περιπτώσεις στην πράξη, η υπάρχουσα γνώση είναι μη πλήρης ή ανακριβής. Σε αυτές τις περιπτώσεις πιθανοτικές συναρτήσεις χρησιμοποιούνται, επιτρέποντας έτσι στα συστήματα τεχνητής νοημοσύνης να πάρουν την αβεβαιότητα του προβλήματος υπόψιν.

3) ΜΑΘΗΣΗ : Σε ένα απλό μοντέλο μάθησης μηχανής (σχ. 1.27), το περιβάλλον προμηθεύει μερικές πληροφορίες σε ένα στοιχείο μάθησης, το στοιχείο μάθησης τότε χρησιμοποιεί αυτή την πληροφορία για να κάνει βελτιώσεις σε μια βάση γνώσης, και τελικά το στοιχείο απόδοσης χρησιμοποιεί τη βάση γνώσης για να κάνει την εργασία του. Το είδος της πληροφορίας που δίνεται στην μηχανή από το περιβάλλον έχει συνήθως ατέλειες, με αποτέλεσμα το στοιχείο μάθησης να μην ξέρει από πριν πως να συμπληρώσει ή να αγνοήσει ορισμένες μη υπάρχοντες λεπτομέρειες που δεν έχουν σημασία. Η μηχανή γι’ αυτό λειτουργεί μαντεύοντας, και δέχεται επανατροφοδότηση από το στοιχείο απόδοσης. Ο μηχανισμός επανατροφοδότησης επιτρέπει στην μηχανή να υπολογίσει τις υποθέσεις της και να τις επαναλάβει αν είναι απαραίτητο. Η μάθηση της μηχανής εμπεριέχει δύο μάλλον διαφορετικά είδη επεξεργασίας πληροφορίας: *επαγωγική και συμπερασματική*. Στην *επαγωγική επεξεργασία πληροφορίας*, γενικά πλαίσια και κανόνες αποφασίζονται από τα ακατέργαστα δεδομένα και την εμπειρία. Στη *συμπερασματική επεξεργασία πληροφορίας*, από την άλλη μεριά, γενικοί κανόνες χρησιμοποιούνται για να αποφασιστούν συγκεκριμένα γεγονότα. Μάθηση βασισμένη σε ομοιότητα χρησιμοποιεί επαγωγή, ενώ η απόδειξη ενός θεωρήματος είναι συμπέρασμα από γνωστά αξιώματα και άλλα γνωστά θεωρήματα.

Μάθηση βασισμένη σε εξήγηση χρησιμοποιεί και επαγωγή και συμπέρασμα.

Η σημασία της βάσης γνώσης και οι δυσκολίες της μάθησης έχουν οδηγήσει στην ανάπτυξη διάφορων μεθόδων για αύξηση της βάσης γνώσης. Ιδιαίτερα αν υπάρχουν ειδικό σε ένα ορισμένο πεδίο, είναι ευκολότερο να συλλεχτεί η εμπειρία των ειδικών απ’ ό,τι να γίνει προσπάθεια να αντιγραφεί η εμπειρία που οδήγησε στην εμπειρογνωμοσύνη. Αυτή πραγματικά είναι η ιδέα πίσω από τα έμπειρα συστήματα.

Αφού έχουμε οικειοποιηθεί με τις συμβολικές μηχανές τεχνητής νοημοσύνης, πως θα μπορούσαμε να τις συγκρίνουμε με τα νευρωνικά δίκτυα σαν γνωστικά μοντέλα. Για αυτή τη σύγκριση χρησιμοποιούμε τρεις υποδιαίρεσεις: *επίπεδο εξήγησης, στιλ επεξεργασίας και δομή αναπαράστασης*.

1) ΕΠΙΠΕΔΟ ΕΞΗΓΗΣΗΣ : Στην κλασσική τεχνητή νοημοσύνη η έμφαση είναι στο χτίσιμο των συμβολικών αναπαραστάσεων.Τυπικά οι αναπαραστάσεις είναι *ασυνεχείς και αυθαίρετες*: αφηρημένες ιδιότητες και όχι ανάλογες εικόνες. Από την πλευρά της γνώσης, είναι αναντίρρητο ότι η τεχνητή νοημοσύνη υποθέτει την ύπαρξη νοητών αναπαραστάσεων και μοντελοποιεί την γνώση σαν την ακολουθιακή διαδικασία των συμβολικών αναπαραστάσεων.

Οι υποθέσεις που γίνονται στα νευρωνικά δίκτυα σχετικά με το τι αποτελεί μια ικανοποιητική εξήγηση των γνωστικών διαδικασιών, είναι εντελώς διαφορετικές από αυτές της κλασσικής τεχνητής νοημοσύνης. Η έμφαση στα νευρωνικά δίκτυα είναι στην ανάπτυξη *παράλληλα καταναμεμημένων μοντέλων επεξεργασίας (PDP)*. Αυτά τα μοντέλα υποθέτουν ότι η επεξεργασία πληροφορίας λαμβάνει χώρα διαμέσου της αλληλεπίδρασης ενός μεγάλου αριθμού νευρώνων , καθένας από τους οποίους στέλνει διεγερτικά και ανασταλτικά σήματα σε άλλους νευρώνες στο δίκτυο. Επιπλέον τα νευρωνικά δίκτυα δίνουν μεγάλη έμφαση στην νευροβιολογική εξήγηση γνωστικών φαινομένων.

2) ΣΤΙΛ ΕΠΕΞΕΡΓΑΣΙΑΣ: Πιθανόν η έμπνευση για ακολουθιακή επεξεργασία προκύπτει από την ακολουθιακή φύση της φυσικής γλώσσας και την λογική καθώς και εξαιτίας της *μηχανής von Neumann*. Δε θα πρέπει να ξεχνιέται ότι η κλασσική τεχνητή νοημοσύνη γεννήθηκε λίγο μετά την μηχανή von Neumann, στον ίδιο πνευματικό κύκλο.

Από την άλλη μεριά, η παράλληλη επεξεργασία είναι ένα εξέχον χαρακτηριστικό των νευρωνικών δικτύων. Πραγματικά, ο παραλληλισμός είναι όχι μόνο εννοιολογικά ουσιαστικός στην επεξεργασία της πληροφορίας στα νευρωνικά δίκτυα, αλλά είναι και η πηγή της ελαστικότητάς τους. Επιπλέον ο παραλληλισμός μπορεί να είναι μαζικός, το οποίο δίνει στα νευρωνικά δίκτυα μια αξιοσημείωτη ακμαιότητα. Με τον παραλληλισμό να απλώνεται σε πολλούς νευρώνες, συνήθως δεν έχει και πολύ σημασία αν οι καταστάσεις ορισμένων νευρώνων αποκλίνουν από τις αναμενόμενες τιμές. Θορυβώδεις ή μη ολοκληρωμένες εισοδοι μπορούν ακόμα να αναγνωριστούν, ένα δίκτυο με βλάβη μπορεί ακόμα να λειτουργήσει ικανοποιητικά και η μάθηση δεν πρέπει να είναι τέλεια. Το δίκτυο γίνεται ακόμα πιο ακμαίο λόγω της “τραχειάς κωδικοποίησης”, όπου κάθε χαρακτηριστικό απλώνεται σε ορισμένους νευρώνες. Γι’αυτό η παράλληλα καταναμεμημένη επεξεργασία πλησιάζει την ελαστικότητα ενός συνεχούς συστήματος, σε οξεία αντίθεση με την ακαμψία της συμβολικής διακριτής τεχνητής νοημοσύνης.

Ένα άλλο ενδιαφέρον συμπέρασμα του παραλληλισμού είναι η *αυτόματη επεξεργασία της πληροφορίας περιεχομένων*. Η γνώση δεν αναπαρίσταται από δηλωτικές εκφράσεις, αλλά από την ποικίλη δομή και την κατάσταση ενεργοποίησης του νευρωνικού δικτύου. Το απαραίτητο περιεχόμενο για ένα δοσμένο πρόβλημα είναι τότε τίποτα λιγότερο από ολόκληρο το νευρωνικό δίκτυο.Κάθε νευρώνας πιθανώς επηρεάζεται από την ολική ενέργεια όλων των άλλων νευρώνων, με αποτέλεσμα η διαχείριση του περιεχομένου να γίνεται αυτόματα.

Εν συντομία, η ακολουθιακή επεξεργασία πληροφορίας εννοεί την κλασσική τεχνητή νοημοσύνη, ενώ η παράλληλη επεξεργασία πληροφορίας εννοεί τα νευρωνικά δίκτυα. Παρόλα αυτά , τα συμβολικά συστήματα τεχνητής νοημοσύνης μπορούν να πραγματοποιηθούν και παράλληλα επίσης. Για παράδειγμα, η γλώσσα LISP, που έχει

γίνει μια σημαντική γλώσσα που χρησιμοποιείται στην τεχνητή νοημοσύνη, έχει πραγματοποιηθεί σε έναν μαζικά παράλληλο υπολογιστή που καλείται μηχανή σύνδεσης. Για τον ίδιο λόγο, δεν υπάρχει λόγος για τον οποίο ακολουθιακή επεξεργασία δεν θα μπορούσε να πραγματοποιηθεί σε μία νευρωνική αρχιτεκτονική. Στην πραγματικότητα τα νευρωνικά δίκτυα εξομοιώνονται με λογισμικό ακολουθιακά.

3) ΔΟΜΗ ΑΝΑΠΑΡΑΣΤΑΣΗΣ : Με μία γλώσσα που πιστεύεται ότι είναι ένα μοντέλο κλασσικής τεχνητής νοημοσύνης, βρίσκουμε ότι οι συμβολικές αναπαραστάσεις έχουν μια *δήθεν γλωσσολογική μορφή*. Όπως οι εκφράσεις της φυσικής γλώσσας, οι εκφράσεις της κλασσικής τεχνητής νοημοσύνης είναι γενικά πολύπλοκες, φτιαγμένες με ένα συστηματικό τρόπο από απλά σύμβολα. Δοσμένου ενός μικρού συνόλου συμβόλων, νέες εκφράσεις με νόημα μπορούν να σχηματιστούν λόγω της ικανότητας σύνθεσης των συμβολικών εκφράσεων και την αναλογία ανάμεσα στην συντακτική δομή και την σημαντική.

Η φύση και δομή των αναπαραστάσεων είναι παρόλα αυτά , ένα κρίσιμο πρόβλημα για τα νευρωνικά δίκτυα. Πραγματικά αυτό το θέμα έχει γίνει ζήτημα διαμάχης στην βιβλιογραφία ανάμεσα σε θεωρητικούς της κλασσικής τεχνητής νοημοσύνης και των νευρωνικών δικτύων. Το Μάρτιο του 1988 σε μία ειδική έκδοση του περιοδικού Cognition οι Fodor και Pylyshyn κάνουν ορισμένες κριτικές για την υπολογιστική επάρκεια των νευρωνικών δικτύων σε σχέση με τη γνώση και την γλωσσολογία. Υποστηρίζουν ότι τα νευρωνικά δίκτυα είναι λάθος ως προς δύο βασικά θέματα της γνώσης: την φύση των νοητικών αναπαραστάσεων και τη φύση των νοητικών διαδικασιών. Σύμφωνα με αυτούς για κλασσική θεωρία τεχνητής νοημοσύνης αλλά όχι νευρωνικά δίκτυα έχουμε:

Οι νοητικές αναπαραστάσεις χαρακτηριστικά εμφανίζουν μια συνδυαστική συντακτική δομή και μια συνδυαστική σημασιολογία.

Οι νοητικές διαδικασίες είναι χαρακτηριστικά ευαίσθητες στην συνδυαστική δομή των αναπαραστάσεων στις οποίες ενεργούν. Σε ένα νευρωνικό δίκτυο, οι αναπαραστάσεις είναι *κατανεμημένες*. Δεν έπεται παρόλα αυτά ότι οτιδήποτε είναι κατανεμημένο πρέπει να έχει συστατικά, και το να είναι κατανεμημένο είναι πολύ διαφορετικό από το να έχει σημασιολογική ή συντακτική δομή χαρακτηριστικών. Δυστυχώς τα περισσότερα από τα μοντέλα νευρωνικών δικτύων που έχουν προταθεί μέχρι σήμερα για κατανεμημένη δομή αναπαραστάσεων είναι μάλλον ad hoc, λύνουν το πρόβλημα για μία συγκεκριμένη κατηγορία που δεν μπορεί να επεκταθεί εύκολα.

Για να κάνουμε έναν απολογισμό, μπορούμε να περιγράψουμε την *συμβολική τεχνητή νοημοσύνη* σαν την *τυπική διαχείριση μίας γλώσσας αλγορίθμων και δεδομένων αναπαραστάσεων με ένα "από την κορυφή προς τη βάση" τρόπο*. Από την άλλη μεριά μπορούμε να περιγράψουμε τα *νευρωνικά δίκτυα* σαν *παράλληλα κατανεμημένους επεξεργαστές με μια φυσική ικανότητα μάθησης και που συνήθως λειτουργούν με ένα "από κάτω προς τα πάνω" τρόπο*. Για την τεκμηρίωση των γνωστικών εργασιών φαίνεται ότι, από το να ψάξουμε λύσεις βασισμένες σε συμβολική τεχνητή νοημοσύνη ή σε νευρωνικά δίκτυα ξεχωριστά, μία πιθανώς πιο χρήσιμη προσέγγιση είναι να χτίσουμε δομημένα συνδετικά μοντέλα που συνδυάζουν και τα δύο. Κάνοντας αυτό, είμαστε σε θέση να συνδυάσουμε τα επιθυμητά χαρακτηριστικά της προσαρμοστικότητας, ακμαιότητας και ομοιομορφίας που προσφέρονται από τα νευρωνικά δίκτυα με την αναπαράσταση, την ικανότητα εξαγωγής συμπερασμάτων και την καθολικότητα που είναι κληρονομημένα χαρακτηριστικά της συμβολικής τεχνητής νοημοσύνης. Για να είναι ένα τέτοιο υβριδικό σύστημα επιτυχές παρόλα αυτά, είναι σημαντικό να υπολογίσουμε

προσεκτικά της περιοχές στις οποίες κάθε μέθοδος δουλεύει καλύτερα. Τεχνικές κλασσικής τεχνητής νοημοσύνης είναι καλύτερες για επεξεργασία φυσικής γλώσσας, σχεδίαση και ξεχωριστή αιτιολόγηση, ενώ τα νευρωνικά δίκτυα είναι καλύτερα για χαμηλού επιπέδου αισθητικές διαδικασίες, ταίριασμα προτύπων και συσχετιστικές μνήμες. Γι' αυτό συνδυάζοντας αυτές τις δύο διαφορετικές μεθόδους μέσα στο ίδιο σύστημα, μπορούμε να συνδυάσουμε τα οφέλη και της κλασσικής τεχνητής νοημοσύνης και των νευρωνικών δικτύων, και να ανακουφίσουμε τις ξεχωριστές αδυναμίες τους. Ένα τέτοιο παράδειγμα υβριδικής αρχιτεκτονικής δικτύου που χρησιμοποιείται για τα ταξινόμηση και υπολογισμό πιθανότητας περιγράφεται από τον Goodman(1992).

1.12 ΙΣΤΟΡΙΚΕΣ ΣΗΜΕΙΩΣΕΙΣ

Καταλήγουμε αυτή την αναφορά πάνω στα νευρωνικά δίκτυα με μερικές ιστορικές σημειώσεις.

Η μοντέρνα περίοδος των νευρωνικών δικτύων λέγεται ότι άρχισε με την πρωτοποριακή δουλειά των McCulloch και Pitts(1943). Ο πρώτος ήταν ψυχίατρος και ο δεύτερος μεγαλοφυής μαθηματικός. Σύμφωνα με τον Rall(1990) η κλασσική εργασία των McCulloch και Pitts έγινε μέσα σε μια κοινωνία που ασχολούνταν με τους νευρώνες στο πανεπιστήμιο του Σικάγο για πάνω από 5 χρόνια. Αυτή η εργασία περιέγραφε το *λογικό λογισμό των νευρωνικών δικτύων*. Είναι αξιοσημείωτο το ότι ο von Neumann χρησιμοποίησε ιδεατά στοιχεία καθυστέρησης τα οποία είχαν υπολογιστεί από τα ιδεατά στοιχεία νευρώνων των McCulloch και Pitts, για την κατασκευή του EDVAC(Electronic Discrete Variable Automatic Computer), ο οποίος κατάληξε στον ENIAC, τον πρώτο γενικού σκοπού υπολογιστή.

Η επόμενη μεγάλη ανάπτυξη πάνω στα νευρωνικά δίκτυα, ήρθε το 1949 με την έκδοση του βιβλίου του Hebb με τίτλο “The Organization of Behavior”, στο οποίο μια ιδιαίτερη δήλωση ενός *φυσιολογικού κανόνα μάθησης για συναπτικές τροποποιήσεις* έγινε για πρώτη φορά. Πιο συγκεκριμένα ο Hebb πρότεινε ότι η *συνδετικότητα του εγκεφάλου συνεχώς αλλάζει καθώς ο οργανισμός μαθαίνει διάφορες εργασίες, και ότι οι νευρωνικοί συγκεντρωτές δημιουργούνται από τέτοιες αλλαγές*. Επίσης πρότεινε το *διάσημο αίτημα μάθησης* σύμφωνα με το οποίο η *αποτελεσματικότητα μιας σύναψης μεταβλητής ανάμεσα σε δύο νευρώνες αυξάνεται από την επαναλαμβανόμενη ενεργοποίηση του ενός νευρώνα από τον άλλο κατά μήκος της σύναψης*.

Η αναφορά των Rochester, Holland, Habit και Duda(1956) είναι ίσως η πρώτη προσπάθεια για *χρήση υπολογιστή που εξομοιώνει*, για να τεσταριστεί μια καλά σχηματισμένη νευρωνική θεωρία βασισμένη στο αίτημα μάθησης του Hebb. Η εξομοίωση έδειξε ότι χρειαζόταν να προστεθεί παρεμπόδιση ώστε η θεωρία πραγματικά να δουλέψει. Τον ίδιο χρόνο ο Uttley παρουσίασε την αποκαλούμενη “διαρρέουσα ολοκλήρωση” ή “νευρώνας φωτιάς” που αργότερα αναλύθηκε από τον Caianiello.

Το 1952 το βιβλίο του Ashby :”Design for a brain: The Origin of Adaptive Behavior” εκδόθηκε, το οποίο ασχολήθηκε με την *βασική έννοια ότι η προσαρμοζόμενη συμπεριφορά δεν είναι έμφυτη αλλά μαθαίνεται*. Το 1954 ο Minsky έγραψε ένα διδακτορικό με τίτλο “Theory of Neural-Analog Reinforcement Systems and Its Application to the Brain-Model problem” και το 1961 ο ίδιος έγραψε μια εργασία με τίτλο “Steps Toward Artificial Intelligence”. Επίσης το 1954 η ιδέα των

μη γραμμικών προσαρμοζόμενων φίλτρων προτάθηκε από τον Gabor (εφευρέτης του ολογραφήματος). Έχτισε μια τέτοια μηχανή και η μάθηση επιτυγχανόταν με τροφοδότηση δειγμάτων στοχαστικών διαδικασιών στη μηχανή, μαζί με συναρτήσεις αποτελεσμάτων που η μηχανή ανέμενε να βγάλει.

Ένα θέμα ιδιαίτερου ενδιαφέροντος για τα νευρωνικά δίκτυα είναι αυτό της σχεδίασης ενός αξιόπιστου δικτύου με νευρώνες που μπορούν να θεωρηθούν σαν μη αξιόπιστα στοιχεία. Αυτό το σημαντικό πρόβλημα λύθηκε από τον von Neumann(1956) χρησιμοποιώντας την ιδέα του πλεονασμού. 15 χρόνια μετά την έκδοση της εργασίας των McCulloch και Pitt μια νέα προσέγγιση πάνω στο πρόβλημα της αναγνώρισης προτύπων έγινε από τον Rosenblatt (1958) στην εργασία του πάνω στο αισθητήριο (perceptron). Το ιδιαίτερο επίτευγμα του ήταν το αποκαλούμενο *θεώρημα σύγκλισης αισθητηρίου*(perceptron convergence theorem). Το 1960 οι Widrow και Hoff πρότειναν τον *αλγόριθμο ελάχιστου μέσου τετραγώνου* (least mean-square:LMS) και τον χρησιμοποίησαν για να σχηματίσουν το *Adaline*(*adaptive linear element*). Η διαφορά ανάμεσα στο αισθητήριο και στο Adaline βρίσκεται στον τρόπο μάθησης. Ένα από τα από τα πρόσφατα εκπαιδευσιμα νευρωνικά δίκτυα με πολλαπλά στοιχεία είναι η δομή Madaline(Widrow). Κατά την διάρκεια της κλασσικής περιόδου του perceptron ('60) πιστευόταν ότι τα νευρωνικά δίκτυα μπορούσαν να κάνουν τα πάντα. Αλλά τότε εκδόθηκε το βιβλίο των Minsky και Papert που με μαθηματικά απέδειξε ότι υπάρχουν όρια πάνω στο τι μπορεί να υπολογιστεί από τα αισθητήρια. Ένα σημαντικό πρόβλημα πάνω στη σχεδίαση ενός πολυεπίπεδου αισθητηρίου είναι το *πρόβλημα της ανάθεσης εμπιστοσύνης*(credit assignment problem), το οποίο βρήκε την λύση του μόλις την δεκαετία του 1980.

Κατά την δεκαετία του '70 λόγω των προβλημάτων εγκαταλείφθηκε το ενδιαφέρον πάνω στα νευρωνικά δίκτυα. Μια σημαντική ενέργεια την δεκαετία αυτή ήταν οι *χάρτες αυτοοργάνωσης με ανταγωνιστική μάθηση*.

Το 1980 έγιναν πολλές εργασίες πάνω στην θεωρία αλλά και στον σχεδιασμό των νευρωνικών δικτύων. Ο Grossberg (1980) ανέπτυξε μια καινούργια αρχή αυτοοργάνωσης που συνδυάζει φιλτράρισμα από "κάτω προς τα πάνω" και αντίθετη αύξηση σε μικρή μνήμη με από "πάνω προς τα κάτω" ταίριασμα προτύπων και σταθεροποίηση του κώδικα μάθησης. Δινόμενης μιας τέτοιας ικανότητας μία δυναμική κατάσταση καλούμενη υιοθετητική ηχώ λαμβάνει χώρα. Αυτό το φαινόμενο δίνει την βάση για μια νέα κατηγορία νευρωνικών δικτύων γνωστά σαν *adaptive resonance theory*(ART).

Το 1982 ο Hopfield χρησιμοποίησε την ιδέα μια συνάρτησης ενέργειας για να φτιάξει ένα νέο τρόπο κατανόησης του υπολογισμού που γίνεται από τα δίκτυα με συμμετρικές συναπτικές συνδέσεις. Επιπλέον καθιέρωσε τον *ισομορφισμό* ανάμεσα σε τέτοια *περιοδικά δίκτυα* και σε ένα *Ising μοντέλο* που χρησιμοποιείται στην στατιστική. Αυτή η αναλογία άνοιξε τον δρόμο για ένα κατακλυσμό θεωριών για τα νευρωνικά δίκτυα. Αυτή η συγκεκριμένη τάξη νευρωνικών δικτύων με ανατροφοδότηση χρήστηκε ιδιαίτερης προσοχής το '80 και με το χρόνο έγιναν γνωστά σαν δίκτυα Hopfield.

Το 1983 οι Cohen και Grossberg έδωσαν μια νέα αρχή για σχεδίαση μιας *διευθυνσιοδοτούμενης μνήμης*(content-addressable memory) που περιλαμβάνει την έκδοση συνεχούς χρόνου του δικτύου Hopfield σαν μια ιδιαίτερη περίπτωση. Μια ακόμα σημαντική ανάπτυξη το 1982 ήταν η έκδοση της εργασίας του Kohonen πάνω στους *χάρτες αυτοοργάνωσης*, χρησιμοποιώντας *μιας ή δύο διαστάσεων δικτυωτές δομές*. Το 1983 οι Kirkpatrick, Gallat και Vecchi περιέγραψαν μια νέα διαδικασία που λέγεται *εξομοιωμένη απόπτηση* για λύση προβλημάτων συνδυαστικής

βελτιστοποίησης. Η εξομοιωμένη ανόπτηση χρησιμοποιείται στην στατιστική θερμοδυναμική και βασίζεται σε μια απλή τεχνική. Την ίδια χρονιά εκδόθηκε μια εργασία από τους Burto, Sutton και Anderson πάνω στην *ενισχυμένη μάθηση*, η οποία δημιούργησε μεγάλο ενδιαφέρον πάνω στην ενισχυμένη μάθηση και την εφαρμογή της. Το 1984 ο Braitenberg εξέδωσε ένα βιβλίο με τίτλο “Vehicles: Experiments in Synthetic Psychology” το οποίο περιγράφει διάφορες μηχανές με απλή εσωτερική αρχιτεκτονική, και το οποίο ενσωματώνει μερικές σημαντικές αρχές της αυτοοργανούμενης απόδοσης. Το 1986 η ανάπτυξη του αλγορίθμου για πίσω διάδοση (back-propagation algorithm) αναφέρθηκε από τον Rumelhart.

Το 1988 ο Linsker περιέγραψε μια νέα αρχή για την αυτοοργάνωση σε ένα δίκτυο αισθητηρίων. Η αρχή αυτή σχεδιάστηκε ώστε να διατηρεί μέγιστη πληροφορία σχετικά με την με τα πρότυπα ενεργειών, που υπόκεινται σε περιορισμούς όπως συναπτικές συνδέσεις και δυναμικές περιοχές σύναψης. Ο Linsker χρησιμοποίησε αφηρημένες έννοιες πάνω στη θεωρία πληροφοριών ώστε να σχηματίσει τη αρχή της διατήρησης μέγιστης πληροφορίας.

Επίσης το 1988 οι Broomhead και Lowe περιέγραψαν μία διαδικασία για το σχεδιασμό “προς τα εμπρός τροφοδότησης”(feedforward) δικτύων χρησιμοποιώντας *συναρτήσεις ακτινικής βάσης(RBF)*, που είναι μια εναλλαγή των πολυεπίπεδων αισθητηρίων. Το 1989 εκδόθηκε το βιβλίο του Mead με τίτλο “VLSI and Neural Systems”. Αυτό το βιβλίο δίνει μια ασυνήθιστη μίξη περιεχομένων από την νευροβιολογία και την τεχνολογία VLSI. Ίσως περισσότερο από κάθε άλλη έκδοση, η εργασία του Hopfield(1982) και το δίτομο βιβλίο των Rummelhart και McLelland(1986), να ήταν οι πιο σημαίνουσες εκδόσεις υπεύθυνες για την αναζωογόνηση του ενδιαφέροντος για τα νευρωνικά δίκτυα στην δεκαετία του ‘80. Τα νευρωνικά δίκτυα έχουν σίγουρα διανύσει πολύ δρόμο από την εποχή των McCulloch και Pitts. Πραγματικά έχουν εγκαθιδρυθεί σαν ενδοπειθαρχικό αντικείμενο με βαθιές ρίζες στην επιστήμη νευρώνων, στην ψυχολογία, στα μαθηματικά, στις φυσικές επιστήμες και στην μηχανική. Δεν είναι αναγκαίο να πούμε ότι είναι εδώ για να μείνουν και θα συνεχίσουν να αναπτύσσονται σε θεωρία, σχεδιασμό και εφαρμογές.

