

Κεφάλαιο 1. Εισαγωγή στα νευρωνικά δίκτυα

Σκοπός: Στο κεφάλαιο αυτό γίνεται μια πρώτη επαφή με τα νευρωνικά δίκτυα. Παρουσιάζονται οι ιδιαιτερότητες του πεδίου αυτού, που το κάνουν να ξεχωρίζει από τα άλλα παραδοσιακά πεδία των φυσικών επιστημών. Δίδονται οι πρώτες ιδέες του τι είναι ένα νευρωνικό δίκτυο, από τι αποτελείται, και σε τι χρησιμεύει. Ποια είναι η σχέση με τα βιολογικά νευρωνικά δίκτυα των ζώντων οργανισμών. Αναπτύσσονται παραστατικά οι αναλογίες με τους υπολογιστές, οι ομοιότητες και οι διαφορές. Τέλος παρουσιάζονται μερικές επιλεγμένες εφαρμογές.

Προσδοκώμενα αποτελέσματα: Όταν θα έχετε τελειώσει τη μελέτη του κεφαλαίου αυτού θα μπορείτε να:

- περιγράψετε ποιοτικά τι είναι ένα νευρωνικό δίκτυο
- να δείξετε τα στοιχεία από τα οποία εμπνέεται η μελέτη των νευρωνικών δικτύων και τη σχέση που έχουν με τα βιολογικά νευρωνικά δίκτυα
- να παραθέσετε τις ομοιότητες και διαφορές μεταξύ των νευρωνικών δικτύων και των γνωστών υπολογιστών
- να απαριθμήσετε μερικές σύγχρονες εφαρμογές των νευρωνικών δικτύων

Έννοιες κλειδιά: νευρωνικά δίκτυα, νευρώνες, συνάψεις, τεχνητά νευρωνικά δίκτυα - ΤΝΔ (artificial neural networks-ANN), νευρωνικά δίκτυα και παράλληλη επεξεργασία, εκπαίδευση νευρωνικών δικτύων, εποπτευόμενη, μη-εποπτευόμενη, και αυτό-εποπτευόμενη εκπαίδευση

Εισαγωγικές παρατηρήσεις: Στο κεφάλαιο αυτό παρουσιάζουμε τα νευρωνικά δίκτυα, ξεκινώντας από την πιο απλή μορφή τους. Δεν απαιτούνται προηγούμενες γνώσεις για την κατανόησή του. Εξηγείται πώς οι πρώτες ιδέες που προέρχονται από τη βιολογία βοηθούν την ανάπτυξη των τεχνητών νευρωνικών δικτύων που υλοποιούνται στους υπολογιστές. Παρουσιάζεται το πιο απλό δίκτυο ενός νευρώνα και οι πρώτες ιδέες για την εκπαίδευση ενός δικτύου. Παρατίθενται οι διαφορές στον τρόπο λειτουργίας ενός νευρωνικού

δικτύου και του κλασικού υπολογιστή, ως προς τη φιλοσοφία που τα διέπει. Τέλος, δίνεται μια σειρά από σύγχρονες εφαρμογές που φαίνεται άμεσα η χρησιμότητά τους, ποιά πράγματα μπορούν να κάνουν με επιτυχία, και τι αποτελέσματα παίρνουμε από αυτά.

Ενότητα 1.1: Τι είναι τα νευρωνικά δίκτυα

Τα νευρωνικά δίκτυα (neural nets) αποτελούν μια σχετικά νέα περιοχή στις φυσικές επιστήμες, καθ' όσον έχουν γίνει γνωστά και έχουν αναπτυχθεί μόνο κατά τα τελευταία σαράντα περίπου χρόνια [1-7]. Εν τούτοις, η περιοχή αυτή έχει δει μια μεγάλη άνθηση, κρίνοντας από την μεγάλη ανάπτυξη που έχει παρατηρηθεί, από τον αριθμό των επιστημόνων που ασχολούνται με αυτά τα θέματα, και βέβαια από τα πολύ σημαντικά επιτεύγματα, που έχουν συμβάλει στο να γίνουν γνωστά σε ένα ευρύτερο κύκλο. Αποτελούν επομένως ένα θέμα με μεγάλο ενδιαφέρον στις τεχνολογικές επιστήμες. Το κύριο χαρακτηριστικό τους είναι ότι οι πρώτες αρχές και λειτουργίες τους βασίζονται στο νευρικό σύστημα των ζώων οργανισμών (και φυσικά του ανθρώπου), αλλά η μελέτη και η χρήση τους έχει προχωρήσει πολύ πέρα από τους βιολογικούς οργανισμούς, και σήμερα τα νευρωνικά δίκτυα χρησιμοποιούνται για να λύσουν κάθε είδους προβλήματα με ηλεκτρονικό υπολογιστή. Η φιλοσοφία τους όμως είναι διαφορετική από τον τρόπο με τον οποίο δουλεύουν οι κλασικοί υπολογιστές. Η λειτουργία τους προσπαθεί να συνδυάσει τον τρόπο σκέψης του ανθρώπινου εγκεφάλου με τον αφηρημένο μαθηματικό τρόπο σκέψης. Έτσι στα νευρωνικά δίκτυα χρησιμοποιούμε ιδέες όπως, π.χ. ένα δίκτυο μαθαίνει και εκπαιδεύεται, θυμάται ή ξεχνά μια αριθμητική τιμή, κλπ. πράγματα που μέχρι τώρα τα αποδίδουμε μόνο στην ανθρώπινη σκέψη. Αλλά βέβαια μπορούν και χρησιμοποιούν επί πλέον και περίπλοκες μαθηματικές συναρτήσεις και κάθε είδους εργαλεία από την μαθηματική ανάλυση.

Ένα ιδιαίτερο χαρακτηριστικό είναι ότι οι επιστήμονες στην περιοχή των νευρωνικών δικτύων προέρχονται σχεδόν από όλες τις περιοχές των φυσικών επιστημών, όπως την ιατρική, την επιστήμη μηχανικών, την φυσική, την χημεία, τα μαθηματικά, την επιστήμη υπολογιστών, ηλεκτρολογία, κλπ. Αυτό δείχνει ότι για την ανάπτυξη τους απαιτούνται ταυτόχρονα γνώσεις και θέματα από πολλές περιοχές, ενώ το ίδιο ισχύει και για τις τεχνικές και τις μεθόδους που χρησιμοποιούνται. Έτσι καταλαβαίνει κανείς ότι τα

νευρωνικά δίκτυα δίνουν μια νέα πρόκληση στις επιστήμες, καθ' όσον οι νέες γνώσεις που απαιτούνται είναι από τις πιο χρήσιμες στον άνθρωπο, τόσο για την ζωή και την ιατρική, όσο και στην τεχνολογία. Καμία άλλη επιστήμη σήμερα δεν συνδυάζει έτσι γνώσεις από τόσο διαφορετικές περιοχές με τόσο άμεσο τρόπο.

Η έμπνευση για κάθε μορφής νευρωνικό δίκτυο ξεκινά από την βιολογία. Οι ζώντες οργανισμοί, από τους πιο απλούς μέχρι τον άνθρωπο, έχουν ένα νευρικό σύστημα, το οποίο είναι υπεύθυνο για μια πλειάδα από διεργασίες, όπως είναι η επαφή με τον εξωτερικό κόσμο, η μάθηση, η μνήμη, κλπ. Το νευρικό σύστημα των οργανισμών αποτελείται από πολλά νευρωνικά δίκτυα τα οποία είναι εξειδικευμένα στις διεργασίες αυτές. Η κεντρική μονάδα του νευρικού συστήματος είναι, οπωσδήποτε, ο εγκέφαλος, ο οποίος επίσης αποτελείται από νευρωνικά δίκτυα. Κάθε νευρωνικό δίκτυο αποτελείται από ένα μεγάλο αριθμό μονάδων, που λέγονται νευρώνες, ή νευρώνια (neurons). Ο νευρώνας είναι η πιο μικρή ανεξάρτητη μονάδα του δικτύου, όπως π.χ. το άτομο είναι η πιο μικρή μονάδα της ύλης. Οι νευρώνες συνεχώς και ασταμάτητα επεξεργάζονται πληροφορίες, παίρνοντας και στέλνοντας ηλεκτρικά σήματα σε άλλους νευρώνες. Βλέπουμε λοιπόν ότι οι πρώτες γνώσεις μας για τα νευρωνικά δίκτυα προέρχονται από την βιολογία και την ιατρική. Σήμερα διεξάγεται ένας μεγάλος όγκος έρευνας στις δύο αυτές επιστήμες για την καλύτερη κατανόηση των νευρωνικών δικτύων του εγκεφάλου, καθ' όσον είναι προφανές ότι αυτό θα βοηθήσει στο να εξηγήσουμε πώς ακριβώς λειτουργεί ο εγκέφαλος, και τις τόσο περίπλοκες διεργασίες του, όπως πώς σκεπτόμαστε, πώς θυμόμαστε, κλπ. Έτσι λοιπόν, τα νευρωνικά δίκτυα των ζώντων οργανισμών τα ονομάζουμε βιολογικά νευρωνικά δίκτυα, ενθουμούμενοι ότι αυτά είναι και τα πρώτα δίκτυα που μελετήθηκαν, καθ' όσον υπάρχουν σε όλους τους ζώντες οργανισμούς (όχι όμως στα φυτά).

Οι διεργασίες που επιτελούνται από τα βιολογικά νευρωνικά δίκτυα στους ζώντες οργανισμούς είναι πολύ περίπλοκες, αλλά και τόσο χρήσιμες στην καθημερινή ζωή του ανθρώπου. Μερικές από αυτές είναι εργασίες ρουτίνας, και τις οποίες ο ανθρώπινος εγκέφαλος εκτελεί με ελάχιστη ή μηδαμινή προσπάθεια, όπως π.χ. αναγνώριση μιας εικόνας. Το ερώτημα που προκύπτει λοιπόν είναι: Μπορούν οι ηλεκτρονικοί υπολογιστές να κάνουν αυτά που κάνει το ανθρώπινο μυαλό; Η απάντηση είναι γνωστή: Πολλά από τα

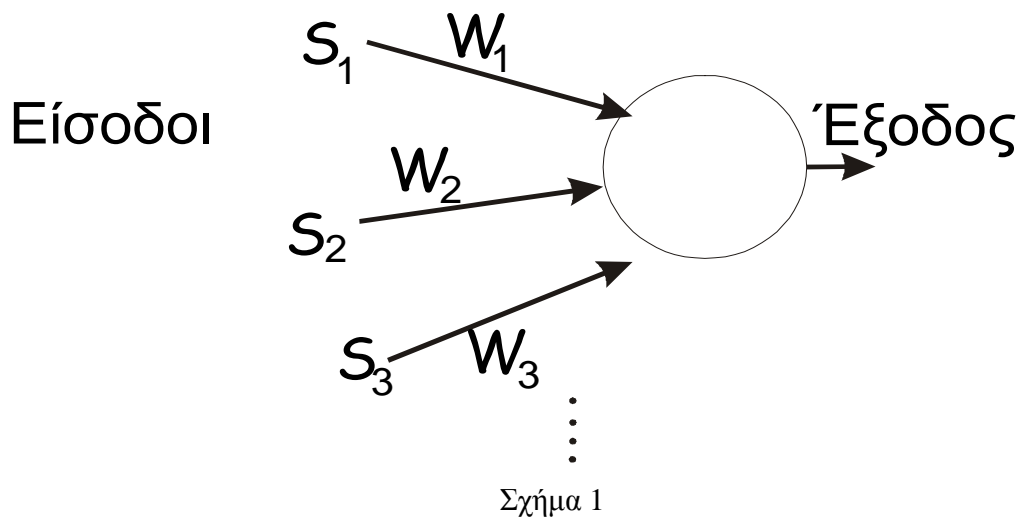
πιό απλά πράγματα, όπως η αναγνώριση φωνής ή εικόνας που το μυαλό κάνει πολύ εύκολα, οι υπολογιστές δεν μπορούν εύκολα να τα κάνουν με επιτυχία. Και βέβαια αυτό δεν οφείλεται στην έλλειψη ταχύτητας, καθ' ότι οι υπολογιστές είναι χιλιάδες φορές γρηγορότεροι από το μυαλό. Ο λόγος είναι ότι η δομή των υπολογιστών είναι πάρα πολύ διαφορετική από την δομή του εγκεφάλου. Το επόμενο λογικό ερώτημα είναι: Θα μπορούσαμε να φτιάξουμε έναν υπολογιστή με τέτοια εσωτερική δομή που να μοιάζει με την δομή του εγκεφάλου, και έτσι να μπορέσουμε να πετύχουμε αυτό που θέλουμε! Αυτό έχει οδηγήσει στο να γίνουν κάποιες πρώτες σκέψεις μήπως είναι δυνατόν να δημιουργηθούν κάποια πρότυπα (μοντέλα) του νευρωνικού συστήματος του ανθρώπου, τα οποία θα περιέχουν όλα τα χαρακτηριστικά που είναι γνωστά μέχρι σήμερα, και τα οποία θα μπορούσαν από μόνα τους να επιτελέσουν τις εργασίες αυτές, με τον ίδιο τρόπο που γίνονται στα βιολογικά νευρωνικά δίκτυα. Τα δίκτυα αυτά ονομάζονται τεχνητά νευρωνικά δίκτυα (artificial neural nets, ANN). Η βασική τους διαφορά από τα βιολογικά δίκτυα είναι ότι τα δίκτυα αυτά παίρνουν γνώσεις (μαθαίνουν) με την εξάσκηση και την εμπειρία, όπως ακριβώς και οι άνθρωποι, αλλά διαφέρουν στο ότι δεν ακολουθούν ορισμένους προκαθορισμένους κανόνες, που είναι χαρακτηριστικό των υπολογιστών. Υπάρχει σήμερα ένας μεγάλος όγκος έρευνας στην περιοχή αυτή, καθ' όσον και εδώ είναι προφανές πόσο χρήσιμο θα ήταν να μπορεί κάτι το άψυχο να επιτελεί εργασίες που μέχρι σήμερα μόνο ο άνθρωπος μπορούσε να κάνει, είτε αυτό είναι μια μηχανή, είτε ένα πρόγραμμα ηλεκτρονικού υπολογιστή. Τα τελευταία τριάντα χρόνια λοιπόν, τα τεχνητά νευρωνικά δίκτυα έχουν δει μεγάλη άνθηση και πρόοδο.

Ο κυρίαρχος σκοπός της λειτουργίας ενός τεχνητού νευρωνικού δικτύου είναι να μπορεί να επιτελεί από μόνο του ορισμένες διεργασίες, π.χ. να αναγνωρίζει εικόνες, αφού όμως προηγουμένως εκπαιδευθεί κατάλληλα. Κάθε δίκτυο δέχεται ορισμένες εισόδους και δίδει ορισμένες εξόδους (input-output). Η εκπαίδευση γίνεται με το να παρουσιάσουμε μια ομάδα από πρότυπα στο δίκτυο, αντιπροσωπευτικά ή παρόμοια με αυτά που θέλουμε να μάθει το δίκτυο. Αυτό σημαίνει ότι δίνουμε στο δίκτυο ως εισόδους κάποια πρότυπα για τα οποία ξέρουμε ποια πρέπει να είναι η έξοδος στο δίκτυο, ξέρουμε δηλ. ποιος είναι ο στόχος, τι πρέπει να δίνει το δίκτυο ως απάντηση στα πρότυπα που του παρουσιάζουμε. Ουσιαστικά είναι σαν να δίνουμε στο δίκτυο την ερώτηση και την

απάντηση που αντιστοιχεί. Το δίκτυο με τα δεδομένα αυτά τροποποιεί την εσωτερική του δομή ώστε να κάνει την ίδια αντιστοιχία που του δώσαμε εμείς. Ακολουθως, αφού βρει την σωστή εσωτερική δομή, τότε θα μπορεί να λύνει και άλλα ανάλογα προβλήματα τα οποία δεν τα έχει δει προηγουμένως, δηλ. δεν έχει εκπαιδευθεί στα πρότυπα των προβλημάτων αυτών. Οποσδήποτε όμως, τα προβλήματα αυτά θα πρέπει να είναι της ίδιας φύσης και των ίδιων χαρακτηριστικών όπως αυτά της εκπαίδευσης και όχι διαφορετικά. Ως σκέψη λοιπόν μια τέτοια διαδικασία είναι πολύ φιλόδοξη. Αυτός είναι ο πιο συνηθισμένος τρόπος εκπαίδευσης, αλλά θα δούμε στο επόμενο τμήμα ότι υπάρχουν διάφορες παραλλαγές ως προς τα δεδομένα που παρουσιάζονται στο δίκτυο όταν εκπαιδεύεται.

Ενότητα 1.2: Πως εκπαιδεύουμε ένα νευρωνικό δίκτυο

Ενα νευρωνικό δίκτυο αποτελείται από ένα αριθμό στοιχείων, τους νευρώνες. Κάθε νευρώνας έχει ένα αριθμό σημάτων που έρχονται ως είσοδος σ' αυτόν, έχει μερικές πιθανές καταστάσεις στις οποίες μπορεί να βρεθεί, και τέλος έχει μία μόνον έξοδο, η οποία είναι συνάρτηση των σημάτων εισόδου, βλέπε Σχήμα 1. Κάθε είσοδος έχει μία δική της τιμή βάρους, και η οποία υποδηλώνει πόσο στενά είναι συνδεδεμένοι οι δύο νευρώνες που συνδέονται με το βάρος αυτό. Η τιμή αυτή συνήθως κυμαίνεται στο διάστημα από -1 ως 1. Η σημασία του βάρους είναι όπως ακριβώς είναι και ο χημικός δεσμός ανάμεσα σε δύο άτομα που απαρτίζουν ένα μόριο. Ο δεσμός μας δείχνει πόσο δυνατά είναι συνδεδεμένα τα δύο άτομα του μορίου. Έτσι και το βάρος μας λέγει πως ακριβώς συνδέονται οι δύο νευρώνες τους οποίους συνδέει.



Σχήμα 1
Ένας νευρώνας (ο κύκλος) με πολλές εισόδους (s_1, s_2, s_3, \dots), αντίστοιχα βάρη (w_1, w_2, w_3, \dots), και μία έξοδο.

Όταν ένας νευρώνας ενεργοποιείται, υπολογίζει μία συνάρτηση από όλα τα δεδομένα που έχει, και συγκρίνει την τιμή της συνάρτησης αυτής με μια τιμή κατωφλίου η οποία είναι χαρακτηριστική για τον νευρώνα αυτόν. Αν η τιμή της συνάρτησης είναι μεγαλύτερη από την τιμή κατωφλίου, τότε ο νευρώνας υπολογίζει την έξοδο, την οποία προωθεί ως είσοδο στον επόμενο (ή στους επόμενους) νευρώνα (νευρώνες). Κατά την διάρκεια της εκπαίδευσης το μόνο πράγμα που αλλάζει είναι η τιμές των βαρών των συνδέσεων των νευρώνων. Οι αλλαγές στις τιμές των βαρών δεν γίνεται πάντα με τον ίδιο τρόπο, αλλά εξαρτάται σημαντικά από την μέθοδο που χρησιμοποιούμε. Στα επόμενα κεφάλαια θα δούμε διάφορες τέτοιες μεθόδους.

Οι αλλαγές αυτές στα βάρη γίνονται με ένα από τους εξής τρεις τρόπους: με εποπτευόμενο τρόπο, με μη-εποπτευόμενο τρόπο, και τέλος με αυτο-εποπτευόμενο τρόπο. Η εποπτευόμενη μάθηση συμβαίνει όταν ξεκινούμε με τυχαίες τιμές στις τιμές των βαρών, και δίνουμε τις τιμές των εισόδων και των στόχων που πρέπει να μάθει το δίκτυο. Κατά την διαδικασία εκπαίδευσης το δίκτυο αλλάζει τις τιμές των βαρών διορθώνοντας

αυτές ανάλογα με το σφάλμα που παίρνουμε (διαφορά από τον στόχο). Στην μη-εποπτευόμενη εκπαίδευση απλώς δίνουμε την πληροφορία στο δίκτυο, χωρίς να γίνεται κανένας έλεγχος. Στην αυτο-εποπτευόμενη εκπαίδευση το δίκτυο αυτο-ελέγχει τον εαυτό του και διορθώνει τα σφάλματα στα δεδομένα με ένα μηχανισμό ανάδρασης (feedback). Σε όλες τις περιπτώσεις όταν το δίκτυο σταματάει να αλλάζει τις τιμές των βαρών, τότε θεωρούμε ότι η εκπαίδευση έχει επιτευχθεί. Αυτό συμβαίνει επειδή το λάθος στην έξοδο γίνεται μηδέν ή είναι πολύ κοντά (τείνει) στο μηδέν.

Ενότητα 1.3: Τα νευρωνικά δίκτυα και οι υπολογιστές

Το Σχήμα 1 δίνει το πιο απλό νευρωνικό δίκτυο που μπορεί να υπάρξει, δηλαδή αποτελείται από έναν μόνο νευρώνα. Πιο περίπλοκα νευρωνικά δίκτυα δημιουργούνται από πολλούς νευρώνες οι οποίοι συνδέονται μεταξύ τους με μια συγκεκριμένη δομή. Καθ' όσον, η δομή τέτοιων δικτύων μπορεί να είναι πολύ περίπλοκη, ομιλούμε πλέον για αρχιτεκτονική δικτύων, πράγμα που αποτελεί ένα από τα καίρια θέματα των τεχνητών νευρωνικών δικτύων. Η αρχιτεκτονική των νευρωνικών δικτύων είναι πολύ διαφορετική από αυτήν των παραδοσιακών υπολογιστών που περιέχουν έναν επεξεργαστή. Οι γνωστοί υπολογιστές δουλεύουν σειριακά, σύμφωνα με τις πρώτες ιδέες του von Neumann, και έχουν την ικανότητα να επιτελούν μερικές εκατοντάδες εντολών, που είναι πολύ γνωστές, όπως είναι οι αριθμητικές πράξεις, κλπ. Στην διαδικασία εκτέλεσης των εντολών ακολουθούν πιστά ένα εσωτερικό ρολόι.

Από τη φύση τους τα νευρωνικά δίκτυα δεν λειτουργούν σειριακά, αλλά με τρόπο που μοιάζει πιο πολύ σε παράλληλο τρόπο λειτουργίας, διότι μία εργασία μοιράζεται στα διάφορα τμήματα του δικτύου, στους νευρώνες, κλπ. Έτσι λέγουμε ότι τα νευρωνικά δίκτυα είναι συστήματα “παράλληλων κατανεμημένων διεργασιών” (“parallel distributed processing”). Αυτό μας παρέχει μεγάλες ταχύτητες, διότι είναι σαν να έχουμε ταυτόχρονα πολλούς επεξεργαστές στη διάθεσή μας. Αλλ' όμως η αρχιτεκτονική των νευρωνικών δικτύων διαφέρει από αυτήν των παραλλήλων επεξεργασιών, καθ' ότι οι απλοί επεξεργαστές των νευρωνικών δικτύων (δηλ. οι νευρώνες) έχουν μεγάλο αριθμό διασυνδέσεων, που συνολικά είναι πολύ μεγαλύτερος από τον αριθμό των νευρώνων. Αντίθετα, στους παράλληλους υπολογιστές, οι επεξεργαστές είναι συνήθως περισσότεροι

από τις διασυνδέσεις μεταξύ τους, και ως προς την πολυπλοκότητα τους ακολουθούν την μηχανή von Neumann. Τα νευρωνικά δίκτυα διαφέρουν από αυτό, διότι οι μονάδες τους είναι πολύ πιο απλές, και επιτελούν πολύ απλούστερες λειτουργίες, δηλ. ξέρουν να αθροίζουν τα σήματα εισόδου, και να τροποποιούν τα βάρη των διασυνδέσεων. Επίσης, οι νευρώνες λειτουργούν ανεξάρτητα ο ένας από τον άλλο και δεν χρειάζονται συγχρονισμό. Αυτό δίνει στα νευρωνικά δίκτυα την ευρωστία και ανοχή σε σφάλματα (βλέπε παρακάτω).

Οι πληροφορίες που αποθηκεύονται σε ένα νευρωνικό δίκτυο μοιράζονται σε ένα μεγάλο αριθμό μονάδων, δηλ. σε πολλούς νευρώνες. Αντίθετα, όταν αποθηκεύουμε στοιχεία στην μνήμη του υπολογιστή, κάθε πληροφορία, σε δυαδική μορφή τοποθετείται σε μια συγκεκριμένη τοποθεσία.

Τελικά, όταν ένα νευρωνικό δίκτυο λύνει ένα πρόβλημα με επιτυχία, παρ' όλο που καταλαβαίνουμε την μαθηματική διαδικασία που ακολουθείται (την οποία εμείς σχεδιάσαμε), εν τούτοις δεν καταλαβαίνουμε γιατί (ή πως) λύνεται το πρόβλημα. Το νευρωνικό δίκτυο δεν "σπάζει" το πρόβλημα σε πολλά μικρά λογικά κομμάτια, αλλά το λύνει με μία "ολιστική" μέθοδο, πράγμα που είναι δύσκολο για το ανθρώπινο μυαλό να το κατανοήσει με απλή λογική. Βέβαια η λύση ελέγχεται εύκολα ότι είναι η σωστή, και έτσι χρησιμοποιείται η τεχνική αυτή με επιτυχία.

Μια άλλη νέα ιδιότητα στα νευρωνικά δίκτυα είναι αυτή της ανοχής σφάλματος. Αυτό σημαίνει ότι αν ένα μικρό τμήμα του δικτύου χαλάσει, το υπόλοιπο δίκτυο συνεχίζει να λειτουργεί, έστω και με ένα μικρό σφάλμα. Αν το δούμε με άλλο τρόπο, σημαίνει ότι αν τα δεδομένα ενός προβλήματος σε ένα μικρό μέρος τους, είναι εσφαλμένα, το δίκτυο δίδει την σωστή απάντηση, και πάλι όμως με ένα μικρό σφάλμα. Είναι γνωστό ότι σε όλα τα παραπάνω οι υπολογιστές δουλεύουν τελείως διαφορετικά. Αν, π.χ. από λάθος σε ένα υπολογιστικό πρόγραμμα ζητήσουμε να γίνει μια διαίρεση μιας ποσότητας δια του μηδενός, τότε ο υπολογιστής σταματά αμέσως την εκτέλεση του προγράμματος, και δίνει μήνυμα σφάλματος, έστω και αν δεν υπάρχει κανένα άλλο σφάλμα στο πρόγραμμα. Αντίθετα ένα νευρωνικό δίκτυο καταλαβαίνει ότι μια τέτοια διαίρεση είναι αδύνατη, την ξεπερνά με κάποιο σφάλμα στο τελικό αποτέλεσμα, και συνεχίζει την λύση του προβλήματος. Βλέπουμε λοιπόν ότι στα νευρωνικά δίκτυα έχουμε κάποια ανοχή στα

σφάλματα. Είναι φυσικό να ρωτήσουμε βέβαια πόσο μεγάλη μπορεί να είναι η ανοχή αυτή. Η απάντηση είναι ότι δεν μπορούμε να προσδιορίσουμε ένα γενικό ποσοστό ανοχής σφάλματος, αλλά τυπικές τιμές σε διάφορα προβλήματα που παρουσιάζονται είναι της τάξης του 10-15%. Όλα όμως εξαρτώνται από το συγκεκριμένο πρόβλημα, και φυσικά υπάρχουν διακυμάνσεις στα νούμερα αυτά.

Το χαρακτηριστικό της ανοχής του σφάλματος στα νευρωνικά δίκτυα είναι μια ιδέα που δεν την συναντάμε σε άλλες συνήθειες υπολογιστικές τεχνικές. Μερικές φορές το στοιχείο αυτό είναι επιθυμητό και λύνει το πρόβλημα μας σχετικά εύκολα, ενώ με άλλες μεθόδους μπορεί να είναι πολύ χρονοβόρο. Είναι ιδιαίτερα χρήσιμο όταν δεν μας ενδιαφέρει η απόλυτη ακρίβεια, αλλά μια προσεγγιστική λύση μπορεί να αρκεί. Αυτό όμως δεν συμβαίνει πάντα, και δεν μπορούμε να πούμε ότι με τα νευρωνικά δίκτυα μπορούμε να λύσουμε όλα τα προβλήματα που μέχρι σήμερα είναι άλυτα. Σε μερικά προβλήματα η χρήση τους δεν συνίσταται καθόλου. Στα επόμενα κεφάλαια θα δούμε λεπτομερώς τις απαιτήσεις που υπάρχουν σε ένα πρόβλημα για να μπορεί ένα νευρωνικό δίκτυο να το αντιμετωπίσει με επιτυχία.

Πίνακας 1: Ομοιότητες και διαφορές μεταξύ των νευρωνικών δικτύων και του υπολογιστή με τη φιλοσοφία του von Neumann.

| | Νευρωνικά Δίκτυα | Υπολογιστής |
|----|---|--|
| 1. | Εργάζονται με σύγχρονο τρόπο λειτουργίας | Εργάζονται με ασύγχρονο τρόπο λειτουργίας |
| 2. | Παράλληλη επεξεργασία | Σειριακή επεξεργασία |
| 3. | Εκπαιδεύονται με παραδείγματα αλλάζοντας τα βάρη των συνδέσεών τους | Προγραμματίζονται με εντολές λογικού χαρακτήρα (if-then) |
| 4. | Η μνήμη, τα όικτυα και οι μονάδες λειτουργίας συνυπάρχουν | Η μνήμη και επεξεργασία πληροφορίας χωρίζονται |
| 5. | Ανοχή στα σφάλματα | Καμία ανοχή στα σφάλματα |
| 6. | Αυτο-οργάνωση κατά τη διαδικασία της | Εξαρτάται εξ ολοκλήρου από το |

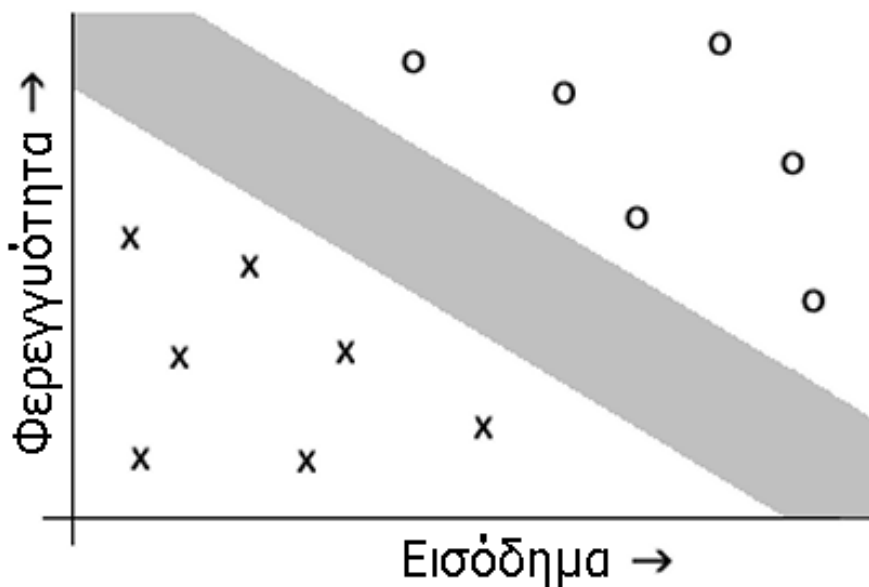
| | | |
|----|--|--|
| | εκπαίδευσης | προσφερόμενο λογισμικό |
| 7. | Η πληροφορία αποθηκεύεται στα βάρη των συνδέσεων | Η πληροφορία αποθηκεύεται σε addressed memory location |
| 8. | Ο χρόνος ενός κύκλου είναι της τάξης του msec | Ο χρόνος ενός κύκλου είναι της τάξης του nsec |

Ενότητα 1.4: Σύγχρονες εφαρμογές των νευρωνικών δικτύων

Όλες οι εφαρμογές των νευρωνικών δικτύων έχουν προκύψει τα τελευταία λίγα χρόνια, και μερικές από αυτές ήδη βρίσκονται ως έτοιμα προϊόντα στην αγορά και χρησιμοποιούνται ευρέως. Είναι βέβαιο ότι τα επόμενα χρόνια ένας πολύ μεγαλύτερος αριθμός θα ακολουθήσει, καθ' όσον το πεδίο αυτό βρίσκεται σε νηπιακή ηλικία. Οι εφαρμογές αυτές περιλαμβάνουν αναγνώριση προτύπων, υπολογισμό συναρτήσεων, βελτιστοποίηση, πρόβλεψη, αυτόματο έλεγχο και άλλα. Θα περιγράψουμε εδώ, μερικές από τις εφαρμογές αυτές:

Στις τραπεζικές εργασίες μια δύσκολη απόφαση είναι να υπολογισθεί ο παράγοντας επικινδυνότητας σε μια αίτηση για ένα στεγαστικό δάνειο. Από τα δεκάδες στοιχεία που περιέχει μια αίτηση, η τράπεζα θέλει να ξέρει τι πιθανότητα υπάρχει ο πελάτης να αδυνατεί να κάνει τις πληρωμές του συμβολαίου. Όπως φαίνεται στο σχήμα 1, τα στοιχεία που παίζουν ρόλο στις αποφάσεις αυτές είναι το εισόδημα και η φερεγγυότητα του δανειολήπτη. Στην περιοχή των "ο", η πιθανότητα να πληρωθεί το δάνειο κανονικά είναι πολύ μεγάλη, ενώ ακριβώς το αντίθετο συμβαίνει στην περιοχή των "x". Ανάμεσα όμως στις δύο περιοχές υπάρχει μια γκριζα περιοχή, στην οποία οι αποφάσεις είναι πολύ δύσκολες. Ένα πρόγραμμα νευρωνικού δικτύου που λέγεται "Νέστωρ" εκπαιδεύεται σε μερικές χιλιάδες αιτήσεις, στις οποίες οι μισές εγκρίθηκαν και οι μισές απορρίφθηκαν από την τράπεζα, με απόφαση των υπαλλήλων της. Συγκρίνοντας με τα πραγματικά δεδομένα, το σύστημα ψάχνει να βρει στοιχεία και να αποφασίσει τι αποτελεί παράγοντα μεγάλης επικινδυνότητας. Τελικά, παίρνει μια απόφαση η οποία έχει μεγαλύτερο ποσοστό επιτυχίας από άλλες μεθόδους. Είναι ιδιαίτερα επιτυχής στην γκριζα περιοχή του σχήματος 1, όπου οι άλλες μέθοδοι δεν δουλεύουν με επιτυχία.. Ένα γνωστό πρόγραμμα

στην εφαρμογή αυτή λέγεται “Nestor” και έχει χρησιμοποιηθεί αρκετά τα τελευταία χρόνια [8,9].



Σχήμα 2

Γράφημα από στοιχεία βάσης δεδομένων που αναφέρεται σε δάνεια, από τα οποία άλλα έχουν αποπληρωθεί σύμφωνα με τους όρους του συμβολαίου και μερικά όχι.

Μια άλλη εφαρμογή είναι η δημιουργία φίλτρου που τοποθετείται σε τηλεπικοινωνιακές γραμμές, όπως πχ οι τηλεφωνικές γραμμές, και το οποίο "καθαρίζει" την γραμμή από τα προβλήματα και τους ήχους και περιορίζει τα σφάλματα κατά τη μετάδοση. Το πρώτο τέτοιο φίλτρο επινοήθηκε από τον B. Widrow, ονομάζεται Adaline και χρησιμοποιείται πάνω από 30 χρόνια με επιτυχία. Είναι από τις πιο παλιές εφαρμογές των νευρωνικών δικτύων [10].

Στην χημική ανάλυση χρησιμοποιούνται νευρωνικά δίκτυα εκεί όπου πρέπει να ληφθούν γρήγορες αποφάσεις και δεν υπάρχει χρόνος για να γίνουν αναλύσεις στο εργαστήριο. Η εταιρία Science Application International (SAIC), έχει δημιουργήσει μια συσκευή θερμικής ανάλυσης νετρονίων (thermal neutron analysis, TNA) που ελέγχεται από ένα νευρωνικό δίκτυο και ανακαλύπτει αντικείμενα τα οποία περιέχουν εκρηκτικά με το να αναλύσει το σήμα εκπομπής ακτίνων γάμα [11]. Η επιτυχία του συνίσταται στο ότι

μπορεί να ξεχωρίσει την προέλευση των στοιχείων, και έτσι μπορεί να καταλάβει και να ξεχωρίσει το σήμα από το άζωτο σε μία βόμβα ή σε ένα γιαούρτι. Σε ένα αεροδρόμιο η ταχύτητα της ανάλυσης είναι 5 δευτερόλεπτα ανά αντικείμενο. Έχει επιτυχία περί το 90%, πράγμα που σημαίνει ότι το ένα στα δέκα αντικείμενα πρέπει να εξετάζεται από υπάλληλο.

Μια άλλη πολύ χρήσιμη εφαρμογή των νευρωνικών δικτύων είναι στην αναγνώριση εικόνων, κειμένων και γενικά προτύπων (pattern recognition). Η εφαρμογή αυτή περιλαμβάνει πάρα πολλές δραστηριότητες, από τις πλέον επιτυχείς των νευρωνικών δικτύων, αλλά εδώ θα αναφέρουμε μόνον ένα απλό πρόγραμμα που σήμερα χρησιμοποιείται κατά κόρον στην επεξεργασία κειμένων. Το πρόγραμμα αυτό λέγεται "OmniPage", το ανέπτυξε η εταιρία Caere το 1994 και υλοποιείται σε ένα απλό PC. Το πρόγραμμα και το προϊόν περιγράφονται λεπτομερώς στην ηλεκτρονική διεύθυνση της εταιρίας Caere, η οποία είναι η εξής : <http://www.caere.com> . Το πρόγραμμα διαβάζει τυπωμένα κείμενα με σαρωτή (scanner) και τα μετατρέπει σε χαρακτήρες ascii. Μάλιστα το πρόγραμμα αυτό δουλεύει, ικανοποιητικά έστω και αν τα γράμματα είναι μερικώς καταστραμμένα, όπως π.χ. από σελίδες fax. Φυσικά, ο αριθμός των εφαρμογών που λειτουργούν σήμερα και βασίζονται σε νευρωνικά δίκτυα είναι πολύ μεγαλύτερος από αυτές που αναφέρονται παραπάνω, οι οποίες είναι μόνον ενδεικτικές των δραστηριοτήτων στην περιοχή αυτή. Αρκετές άλλες θα αναφερθούν στα επόμενα κεφάλαια.

Σύνοψη: Τα τεχνητά νευρωνικά δίκτυα (ΤΝΔ) είναι προγράμματα σε ηλεκτρονικό υπολογιστή, που βασίζονται στο πρότυπο του ανθρώπινου εγκεφάλου. Δεν περιέχουν όλες τις λεπτομέρειες της δομής και λειτουργίας του εγκεφάλου, τα οποία εξ άλλου δεν είναι γνωστά ακόμα και σήμερα. Χρησιμοποιούν μόνο την κεντρική ιδέα της λογικής λειτουργίας του, ξεκινώντας από μία συλλογή μονάδων, που είναι αντίστοιχες προς τους νευρώνες, και προσπαθούν να επιτελέσουν τις ανάλογες διεργασίες. Τελικά όμως τα βιολογικά και τα τεχνητά δίκτυα διαφέρουν πάρα πολύ ως προς την αρχιτεκτονική και τις ιδιότητες τους. Αναπτύχθηκαν την τελευταία πενήνταετία και έχουν πετύχει αρκετά εντυπωσιακά αποτελέσματα, αλλά έχει φανεί επίσης ότι έχουν και αρκετούς περιορισμούς. Οι περιορισμοί δημιουργούνται συχνά όταν το μέγεθος και η

πολυπλοκότητα του συστήματος αυξάνουν. Είναι ιδιαίτερα ικανά σε συνδυαστικά προβλήματα και σε γενικοποιήσεις. Αντίθετα, δεν είναι ικανά σε προβλήματα λογικής και σε ακριβείς υπολογισμούς. Υπάρχουν πάρα πολλά μοντέλα δικτύων με διαφορετική φιλοσοφία και τρόπο λειτουργίας, και πολλές και διάφορες εφαρμογές.

Βιβλιογραφία:

1. J. A. Anderson, An Introduction to Neural Networks, MIT Press, Cambridge (1995).
2. S. Haykin, Neural Networks: A Comprehensive Foundation, Second Edition, Prentice Hall, Upper Saddle Point (1999).
3. N. K. Bose and P. Liang, Neural Networks Fundamentals with Graphs, Algorithms and Applications, McGraw-Hill, New York (1996).
4. R. Lippmann, An Introduction to Computing with Neural Networks, IEEE ASSP Magazine, 4-22 (1987).
5. K. K. Obermeier and J. J. Barron, Time to get fired up, Byte 217 (1987).
6. R. Hecht-Nielsen, Neurocomputing:picking the human brain, IEEE Spectrum 37(1988).
7. T. Kohonen, Adaptive, associative, and self-organizing functions in neural computing, Applied Optics, **26**,4910(1987).
8. P.D. Wasserman, Advanced Methods in Neural Computing, Van Nostrand Reinhold (New York), 1993.
9. M. McCord Nelson and W.T. Illingworth, A practical guide to Neural Nets, Addison-Wesley (Reading, Mass), 1991.
10. B. Widrow and M. E. Hoff, Adaptive Switching Circuits, 1960 WESCON Convention, Record Part 4, pp. 96-104; Human Neurobiology, **4**,229(1985).
11. R.C. Johnson, “Neural Nose to Sniff Out Explosives at JFK Airport”, Electronic Engineering Times **536**,1(1989).