

**ΑΚΑΔΗΜΙΑ ΕΜΠΟΡΙΚΟΥ ΝΑΥΤΙΚΟΥ  
ΜΑΚΕΔΟΝΙΑΣ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**ΘΕΜΑ : ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ  
ΕΦΑΡΜΟΓΕΣ**

**ΣΠΟΥΔΑΣΤΗΣ : ΔΡΑΚΟΝΤΗΣ ΠΑΝΑΓΙΩΤΗΣ**

**ΕΠΙΒΛΕΠΩΝ  
ΚΑΘΗΓΗΤΗΣ : ΠΑΠΑΣΤΑΜΟΥΛΗΣ Α.**

**ΝΕΑ ΜΗΧΑΝΙΩΝΑ  
2013**

**ΑΚΑΔΗΜΙΑ ΕΜΠΟΡΙΚΟΥ ΝΑΥΤΙΚΟΥ  
ΜΑΚΕΔΟΝΙΑΣ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**ΘΕΜΑ : ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ  
ΕΦΑΡΜΟΓΕΣ**

**ΣΠΟΥΔΑΣΤΗΣ : ΔΡΑΚΟΝΤΗΣ ΠΑΝΑΓΙΩΤΗΣ  
ΑΜ : 4120**

**ΗΜΕΡΟΜΗΝΙΑ ΠΑΡΑΔΟΣΗΣ :**

Βεβαιώνεται η ολοκλήρωση της παραπάνω πτυχιακής εργασίας

Ο καθηγητής

## Περίληψη

Το αντικείμενο των τεχνητών νευρωνικών δικτύων γνωρίζει ραγδαία ανάπτυξη τα τελευταία 35 χρόνια και αποτελεί πλέον ένα ευρύ τομέα που σχετίζεται με το γενικότερο πλαίσιο της τεχνητής νοημοσύνης και των ευφυών συστημάτων. Τα τεχνητά νευρωνικά δίκτυα ουσιαστικά προσομοιώνουν τη βιολογική οργάνωση και λειτουργία των βιολογικών νευρώνων και βασικό τους πλεονέκτημα είναι η ευπλαστικότητα.

Στην εργασία αυτή γίνεται μία σύντομη ιστορική αναδρομή στην ανακάλυψη και στην ανάπτυξη των τεχνητών νευρωνικών δικτύων, ενώ ακολούθως περιγράφεται τι ακριβώς είναι ένα τεχνητό νευρωνικό δίκτυο, από τι αποτελείται, ποια είναι η βασική δομή του και παρουσιάζονται τα πιο αναγνωρισμένα μοντέλα αισθητήρων του.

Στην συνέχεια ακολουθεί η ανάλυση της δομής ενός τεχνητού νευρωνικού δικτύου και οι μέθοδοι εκπαίδευσης και μάθησης των δικτύων αυτών ανάλογα με την φύση του εκάστοτε προβλήματος. Ακόμη θα εξετάσουμε τα πλεονεκτήματα και τα μειονεκτήματα των τεχνητών νευρωνικών δικτύων για να αναλύσουμε την αξιοπιστία αυτής της τεχνολογίας σε σχέση με τους ηλεκτρονικούς υπολογιστές. Επίσης παρουσιάζονται ορισμένες εφαρμογές των τεχνητών νευρωνικών δικτύων που χρησιμοποιούνται στην καθημερινότητα των ανθρώπων αλλά και εφαρμογές οι οποίες επέλυσαν σημαντικά προβλήματα σε διάφορους τομείς. Τέλος αναφέρονται και αναλύονται μερικές εφαρμογές τεχνητών νευρωνικών δικτύων στον τομέα της ναυτιλίας και στην συνέχεια αναφέρονται συμπεράσματα για τις εφαρμογές αυτές και την εξέλιξη τους.

## **Abstract**

The subject of artificial neural networks has been rapidly advancing for the last 35 years and comprises a wide field relating to the general framework of artificial intelligence and intelligent systems. Artificial neural network essentially simulate the biological arrangement and function of biological neurons and their main advantage is malleability.

This report covers a brief history of the invention and development of artificial neural networks, while it subsequently describes what exactly is an artificial neural network, what it is comprised of, what is its basic structure and presents its most recognized sensor models.

Then, the analysis of the structure of an artificial neural network follows and the training and teaching methods of these networks in accordance with the nature of each problem. Moreover, we will examine the advantages and disadvantages of artificial neural networks in order to determine the reliability of their technology in relation to computers. In addition, certain applications of artificial neural networks in daily life are presented as well as applications which resolved significant problems in various fields. Finally, reference and analysis is made to some applications of artificial neural networks in shipping and conclusions about these applications and their advancement / progress are drawn.

## Πρόλογος

Τα νευρωνικά δίκτυα γνωρίζουν ανάπτυξη τις 3 τελευταίες δεκαετίες στα ευφυή συστήματα και στα αισθητήρια. Στο πρώτο κεφάλαιο θα αναφερθούμε στην ιστορία της εξέλιξης αυτών των συστημάτων.

Στο δεύτερο κεφάλαιο θα κάνουμε μια εισαγωγή για το τι ακριβώς είναι ένα νευρωνικό δίκτυο αναλύοντας πως εμπνεύστηκε η κατασκευή τους.

Στο τρίτο κεφάλαιο θα εξηγήσουμε την λειτουργία των τεχνητών νευρωνικών δικτύων και θα αναλύσουμε την βασική δομή τους παρουσιάζοντας βασικά παραδείγματα.

Στο τέταρτο κεφάλαιο θα αναφερθούμε στον τρόπο εκπαίδευσης των δικτύων με σκοπό την καλύτερη επίλυση προβλημάτων.

Στο πέμπτο κεφάλαιο προχωράμε στην ανάλυση των αισθητηρίων αυτών των συστημάτων συγκρίνοντας τα με τους ηλεκτρονικούς υπολογιστές παραθέτοντας σημαντικές διαφορές και ομοιότητες που υπάρχουν μεταξύ τους.

Στο έκτο κεφάλαιο θα εξετάσουμε εφαρμογές αυτών των δικτύων σε διάφορους επαγγελματικούς τομείς.

Τέλος στο έβδομο κεφάλαιο θα παρουσιάσουμε παραδείγματα της εφαρμογής των δικτύων σε δραστηριότητες και κλάδους της ναυτιλίας τονίζοντας τον σημαντικό ρόλο που διατηρούν τα νευρωνικά δίκτυα σε αυτήν.

# Κεφάλαιο 1

## ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ

Τα νευρωνικά δίκτυα είναι σχετικά μία νέα περιοχή, δεν υπάρχει ουσιαστικά μεγάλη προϊστορία, όπως σε άλλες παραδοσιακές επιστήμες. Ξεκίνησε σε διεθνές επίπεδο μόλις κατά τις τελευταίες δεκαετίες, αλλά η μεγάλη ώθηση σ' αυτά δόθηκε μετά το 1980. Σ' αυτό βοήθησε τόσο η τεράστια ανάπτυξη του υλικού/λογισμικού των Η/Υ όσο και η ανάπτυξη νέων αλγορίθμων εκπαίδευσης. Η ανάπτυξη των νευρωνικών δικτύων πέρασε από πολλές φάσεις και εξελίξεις.

Το πρώτο μοντέλο νευρωνικού δικτύου το οποίο προτείνει ότι οι νευρώνες είναι η βασική μονάδα του δικτύου παρουσιάστηκε το 1943 από τους McCulloch και Pitts. Σε μία πρώτη εργασία τους οι ερευνητές αυτοί παρουσίασαν για πρώτη φορά την ιδέα ότι ένα νευρωνικό δίκτυο αποτελείται από μία συλλογή ενός μεγάλου αριθμού νευρώνων και έδειξαν πώς θα μπορούσαν να λειτουργούν οι νευρώνες με τις διασυνδέσεις τους. Αυτή θεωρείται ιστορικά ότι είναι η πρώτη εικόνα ενός νευρωνικού δικτύου. Μάλιστα οι συγγραφείς θεώρησαν ότι οι νευρώνες και οι συνδέσεις τους είναι ένα πρότυπο, ανάλογο ενός ηλεκτρικού κυκλώματος.

Ο McCulloch ήταν νευροφυσιολόγος και ο Pitts ένας 18χρονος πρωτοετής φοιτητής των Μαθηματικών. Οι ίδιοι συγγραφείς προχώρησαν το 1947 σε πιο εξελιγμένο πρότυπο για την αναγνώριση σχημάτων. Το πρότυπο αυτό περιέχει πολλά χαρακτηριστικά από τα μεταγενέστερα πρότυπα. Ο νευρώνας θεωρείται ότι μπορεί να έχει δύο μόνον καταστάσεις. Μπορεί να δέχεται πολλές εισόδους αλλά δίνει μία μόνον έξοδο. Οι έξοδοι από διαφορετικούς νευρώνες δεν επιτρέπεται να ενώνονται, αλλά πρέπει υποχρεωτικά να οδηγούν σε είσοδο άλλου νευρώνα. Οι απολήξεις των νευρώνων είναι δύο ειδών: διεγερτικές και ανασταλτικές. Οι δύο καταστάσεις του νευρώνα είναι ότι είτε πυροδοτεί ή βρίσκεται σε ηρεμία. Η ροή της πληροφορίας μέσα στον νευρώνα ελέγχεται από πύλες, οι οποίες επίσης είναι διεγερτικές ή ανασταλτικές. Όταν ο νευρώνας πυροδοτεί, στέλνει ένα παλμό. Οι λειτουργίες αυτές πάντα γίνονται σε διακριτό χρόνο και υποτίθεται ότι όλοι οι νευρώνες αποκρίνονται ταυτόχρονα, δηλαδή το σύστημα δρα συγχρονισμένα. Η κατάσταση ενός νευρώνα σε χρόνο  $t + 1$  εξαρτάται από την κατάστασή του σε χρόνο  $t$  και από τις εισόδους που εισέρχονται στην χρονική αυτή στιγμή.

Τα δίκτυα McCulloch–Pitts προσπαθούν να εξηγήσουν για πρώτη φορά πως δουλεύει η μνήμη. Θεωρούν ότι ένας πιθανός μηχανισμός μνήμης μπορεί να είναι η ύπαρξη κλειστών διαδρομών του σήματος μέσα στο δίκτυο. Αν δεν υπάρχει καμιά τέτοια διαδρομή και χωρίς νέο εξερχόμενο σήμα, τότε το δίκτυο θα μείνει μόνιμα σε κατάσταση ηρεμίας. Έτσι, μια ίνα ενώνει την έξοδο ενός κυττάρου με το σημείο εισόδου στο ίδιο κύτταρο, δημιουργώντας έναν μηχανισμό ανάδρασης (feedback). Μόλις πυροδοτεί ένα τέτοιο κύτταρο θα συνεχίσει να πυροδοτεί μέχρι να έλθει σήμα από ανασταλτική ίνα. Καθ'όλη τη διάρκεια της λειτουργίας αποστέλλονται παλμοί στην πύλη των κυττάρων και μεταδίδεται το σήμα και η πληροφορία. Αυτός ο κύκλος του σήματος θα μπορούσε να θεωρηθεί ως μία πρώτη ιδέα για ένα μηχανισμό μνήμης. Οι εργασίες αυτές πιθανόν να χάνονταν στην βιβλιογραφία αν δεν τις χρησιμοποιούσε λίγα χρόνια αργότερα ο J. von Neumann ως παράδειγμα για υπολογιστικές μηχανές την δεκαετία που διαδόθηκε ο ηλεκτρονικός υπολογιστής, δηλαδή την δεκαετία του 1950. Τότε έγιναν και οι πρώτες προσπάθειες να αντλήσουμε πληροφορίες από τα βιολογικά δίκτυα και να δημιουργηθούν τα πρώτα τεχνητά δίκτυα.

Ένα άλλο έργο της πρώτης αυτής εποχής που αφήνει ακόμα και σήμερα την επιρροή του είναι το βιβλίο του D. Hebb, «The organization of behavior» (1949), το οποίο εισάγει τον κανόνα μάθησης του Hebb. Το μοντέλο του Hebb έχει ως κεντρική ιδέα τις συνδέσεις μεταξύ μονάδων του συστήματος, δηλαδή τους νευρώνες. Έφτασε στα συμπεράσματα αυτά μετά από σωρεία πειραμάτων νευροφυσιολογίας. Ο κανόνας αυτός λέγει ότι κάθε φορά που το δίκτυο χρησιμοποιεί τις νευρωνικές του συνδέσεις, οι συνδέσεις αυτές ενισχύονται και το δίκτυο πλησιάζει περισσότερο στο να μάθει το πρότυπο το οποίο παρουσιάζεται. Όταν ο νευρώνας  $i$  επανειλημμένα διεγείρει τον νευρώνα  $j$ , τότε συμβαίνει να αναπτύσσεται μια μεταβολική σύνδεση στον ένα ή και στους δύο νευρώνες, έτσι ώστε η απόδοση του φαινομένου (το  $i$  διεγείρει το  $j$ ) να αυξάνεται. Αν  $w_{ij}$  είναι το βάρος της σύνδεσης μεταξύ  $i$  και  $j$ ,  $x_i$  η είσοδος στον νευρώνα  $j$  από τον νευρώνα  $i$ ,  $x_j$  η έξοδος του νευρώνα  $j$ , τότε ισχύει ότι:  $w_{ij}(\text{new}) = w_{ij}(\text{old}) + ax_ix_j$ . Εδώ  $a$  είναι μία θετική σταθερά που λέγεται παράμετρος του ρυθμού εκπαίδευσης. Το νέο λοιπόν βάρος  $w_{ij}$  θα είναι ίσο με το παλαιό ενισχυμένο κατά μία ποσότητα  $ax_ix_j$ . Ο κανόνας αυτός έχει τοπικό χαρακτήρα, ισχύει δηλαδή μόνο για την σύνδεση του νευρώνα  $i$  και  $j$  και όχι για άλλες συνδέσεις του δικτύου. Στην επόμενη

παράγραφο θα αναφέρουμε βασικά μοντέλα αισθητήρων που εφευρέθηκαν στο πέρασμα των χρονών.

Ένα από τα βασικότερα μοντέλα αισθητήρων είναι το μοντέλο perceptron. Το μοντέλο του αισθητήρα (perceptron) παρουσιάστηκε για πρώτη φορά το 1957 από τον F. Rosenblatt, ο οποίος αρχικά έφτιαξε το πρώτο δίκτυο με hardware που μπορούσε να κάνει πολλές και διάφορες διεργασίες. Είναι ένα πολύ απλό μοντέλο (οι λεπτομέρειες θα παρουσιασθούν στα επόμενα κεφάλαια) που έχει μόνο δύο επίπεδα, αυτά της εισόδου και της εξόδου. Το σήμα προχωρά μονοδρομικά από την είσοδο στην έξοδο.

Το μοντέλο αυτό στην αρχή είχε πολλές επιτυχίες, δημιούργησε μεγάλο ενθουσιασμό και μάλιστα ήδη αρχίζει να συζητείται η ιδέα ότι πιθανόν τα νευρωνικά δίκτυα να είναι η ανώτερη τεχνική που λύνει όλα τα προβλήματα που μέχρι τότε παρέμεναν άλυτα. Οι πρώτες λοιπόν επιτυχίες μεγαλοποιήθηκαν, αλλά γρήγορα φάνηκε ότι τα μοντέλα αυτά είχαν πολλούς περιορισμούς. Μια συνολική και εμπειριστατωμένη εικόνα του προτύπου αυτού παρουσιάστηκε το 1969 στο βιβλίο *Perceptrons* των Minsky και Papert. Στο βιβλίο αυτό γίνεται μία συνολική εκτίμηση της χρησιμότητας του προτύπου του αισθητήρα και όλων των διεργασιών για τα οποία είναι χρήσιμο. Αποδεικνύεται με αναλυτικά μαθηματικά ότι υπάρχουν συγκεκριμένοι περιορισμοί στο πρότυπο αυτό. Έτσι, δεν μπορεί να λύσει, λ.χ. το σχετικό πρόβλημα του X-OR (αναφέρεται στο κεφάλαιο 6).

Οι αρχικές προσδοκίες που είχαν δημιουργηθεί ήδη φαίνεται ότι δεν επαληθεύονται και προς το παρόν τα νευρωνικά δίκτυα χάνουν την δημοτικότητα τους, με αποτέλεσμα ο κόσμος να στρέφεται σε μια νέα παρεμφερή περιοχή που τότε άρχισε να γίνεται γνωστή, την Τεχνητή Νοημοσύνη. Την ίδια περίπου εποχή με την ανάπτυξη του μοντέλου του αισθητήρα οι Widrow και Hoff ανέπτυξαν το 1959 δύο νέα μοντέλα, το Adaline και το Madaline, τα οποία όπως είδαμε νωρίτερα ήταν από τα πρώτα μοντέλα που χρησιμοποιήθηκαν επιτυχώς για πρακτικά προβλήματα: Χρησιμοποιήθηκαν ως φίλτρα για να εξαλείψουν την ηχώ σε τηλεφωνικές γραμμές. Τα επόμενα είκοσι χρόνια, μέχρι περίπου το 1980, μικρή μόνο πρόοδος επιτελέστηκε στα νευρωνικά δίκτυα, διότι οι περιορισμοί που αναφέρθηκαν παραπάνω αποθάρρυναν τους περισσότερους στο πεδίο αυτό, το οποίο, όπως φάνηκε λίγο αργότερα, έψαχνε να βρει μία διέξοδο και να κάνει νέα σημαντικά βήματα.



## Κεφάλαιο 2

### ΤΙ ΕΙΝΑΙ ΕΝΑ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ

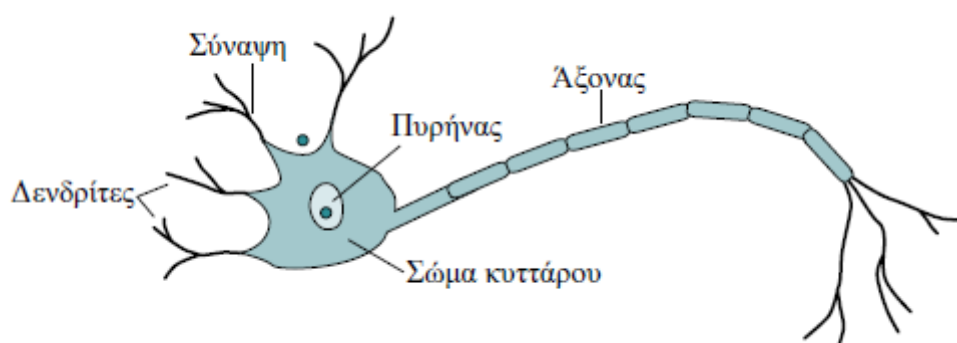
Τα νευρωνικά δίκτυα (neural networks) αποτελούν μια σχετικά νέα περιοχή στις φυσικές επιστήμες, καθόσον έχουν γίνει γνωστά και έχουν αναπτυχθεί σε διεθνές επίπεδο μόνο κατά τις τελευταίες δεκαετίες. Εν τούτοις, η περιοχή αυτή έχει δει μια μεγάλη άνθηση, η οποία διαφαίνεται από την μεγάλη ανάπτυξη που έχει παρατηρηθεί, από τον αριθμό των επιστημόνων που ασχολούνται με αυτά τα θέματα και βέβαια από τα πολύ σημαντικά επιτεύγματα που έχουν συμβάλλει στο να γίνουν τα νευρωνικά δίκτυα 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 και εμπνέονται από το νευρικό σύστημα των ζώντων οργανισμών (και φυσικά του ανθρώπου), αλλά η μελέτη και η χρήση τους έχει προχωρήσει πολύ πέρα από τους βιολογικούς οργανισμούς. Ουσιαστικά δημιουργήθηκε μία νέα περιοχή η οποία έχει αποκοπεί τελείως από την βιολογία 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 τέτοιες ιδέες όπως, π.χ. ένα δίκτυο μαθαίνει και εκπαιδεύεται, θυμάται ή ξεχνά μια αριθμητική τιμή κτλ., πράγματα που μέχρι τώρα τα αποδίδαμε 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, τη Φυσική, τη Χημεία, τα Μαθηματικά, την επιστήμη Υπολογιστών, την Ηλεκτρολογία κτλ. Αυτό δείχνει ότι για την ανάπτυξή 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, καθόσον οι νέες γνώσεις που απαιτούνται είναι από τις πιο χρήσιμες στον άνθρωπο, τόσο για την ζωή και την ιατρική όσο και για την τεχνολογία. Καμία άλλη επιστήμη σήμερα δεν συνδυάζει με τόσο άμεσο τρόπο γνώσεις που προέρχονται από τόσο διαφορετικές περιοχές.

Η έμπνευση για τα νευρωνικά δίκτυα, όπως αναφέρθηκε παραπάνω, 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 κτλ. 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. Στο σχήμα 2.1 παρουσιάζεται η δομή ενός νευρώνα του ανθρώπινου εγκεφάλου.



**Εικόνα 2.1:** Τυπικός νευρώνας ανθρώπινου εγκεφάλου

Βλέπουμε λοιπόν ότι οι πρώτες γνώσεις μας για τα νευρωνικά δίκτυα προέρχονται από την βιολογία και την ιατρική. Σήμερα διεξάγεται ιδιαίτερα μεγάλη έρευνα στις δύο αυτές επιστήμες για την καλύτερη κατανόηση των νευρωνικών δικτύων του εγκεφάλου, καθόσον 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 κτλ.

Οι έννοιες αυτές, παρόλο ότι ακούγονται απλές, εν τούτοις δεν έχουν εξηγηθεί σχεδόν καθόλου μέχρι σήμερα από τους επιστήμονες. Έτσι, λοιπόν, τα νευρωνικά δίκτυα των ζώντων οργανισμών τα ονομάζουμε βιολογικά νευρωνικά δίκτυα, ενθυμούμενοι ότι αυτά είναι και τα πρώτα δίκτυα που μελετήθηκαν, καθόσον 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, τις οποίες ο ανθρώπινος εγκέφαλος εκτελεί με ελάχιστη ή μηδαμινή προσπάθεια, όπως π.χ. η 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 πιο 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, καθώς 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.

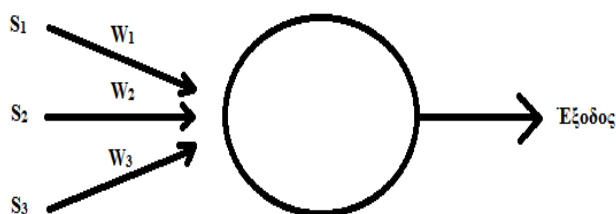
Τα δίκτυα αυτά ονομάζονται τεχνητά νευρωνικά δίκτυα (artificial neural nets, ANN). Η βασική τους διαφορά από τα βιολογικά δίκτυα είναι ότι τα δίκτυα αυτά παίρνουν γνώσεις (μαθαίνουν) με την εξάσκηση και την εμπειρία, όπως ακριβώς και οι άνθρωποι, αλλά διαφέρουν στο ότι δεν ακολουθούν ορισμένους προκαθορισμένους κανόνες, που είναι χαρακτηριστικό των υπολογιστών. Υπάρχει σήμερα ένας μεγάλος όγκος έρευνας στην περιοχή αυτή, καθώς 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. Τις τελευταίες λοιπόν δεκαετίες, στην προσπάθεια να απαντηθούν τα ερωτήματα αυτά, τα τεχνητά νευρωνικά δίκτυα έχουν δει μεγάλη άνθηση και πρόοδο.

## Κεφάλαιο 3

### ΔΟΜΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

Ένα νευρωνικό δίκτυο αποτελείται από ένα αριθμό στοιχείων, τους νευρώνες. Σε κάθε νευρώνα καταφθάνει ένας αριθμός σημάτων, τα οποία έρχονται ως είσοδος σε αυτόν. Ο νευρώνας έχει μερικές πιθανές καταστάσεις στις οποίες μπορεί να βρεθεί η εσωτερική δομή του που δέχεται τα σήματα εισόδου και, τέλος, έχει μία μόνον έξοδο, η οποία είναι συνάρτηση των σημάτων εισόδου. Κάθε σήμα που μεταδίδεται από ένα νευρώνα σε ένα άλλο μέσα στον νευρωνικό δίκτυο συνδέεται με την τιμή βάρους,  $w$ , και η οποία υποδηλώνει πόσο στενά είναι συνδεδεμένοι οι δύο νευρώνες που συνδέονται με το βάρος αυτό. Η τιμή αυτή συνήθως κυμαίνεται σε ένα συγκεκριμένο διάστημα, π.χ. στο διάστημα από  $-1$  ως  $1$ , αλλά αυτό είναι αυθαίρετο και εξαρτάται από το πρόβλημα που προσπαθούμε να λύσουμε. 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 ένα βάρος μας λέγει ακριβώς πόσο σημαντική είναι η συνεισφορά του συγκεκριμένου σήματος στην διαμόρφωση της δομής του δικτύου για τους δύο νευρώνες τους οποίους συνδέει. Όταν το  $w$  είναι μεγάλο (μικρό), τότε η συνεισφορά του σήματος είναι μεγάλη (μικρή).



Σχήμα 3.1: Τυπική δομή νευρώνα νευρωνικού δικτύου

Τώρα θα δούμε πως και με ποια διαδικασία μεταδίδεται το σήμα από νευρώνα σε νευρώνα. Ουσιαστικά όλα τα σήματα που φθάνουν σε ένα νευρώνα αθροίζονται, υπόκεινται σε μία διαδικασία. Παράγεται ως αποτέλεσμα της διαδικασίας μία έξοδος και αυτό είναι το σήμα το οποίο μεταδίδεται περαιτέρω στους επόμενους νευρώνες.

Η θεώρηση αυτή είναι γενική και ισχύει πάντοτε, αυτό όμως που αλλάζει είναι η διαδικασία η οποία δεν είναι πάντα η ίδια. Ας δούμε αρχικά δύο τρόπους με τους οποίους γίνεται η μετάδοση αυτή. Ο πρώτος τρόπος είναι δυαδικός. Στην περίπτωση αυτή ένας νευρώνας μπορεί να βρεθεί σε μία από δύο δυνατές καταστάσεις: να είναι ενεργός ή να είναι αδρανής. Όταν ένας νευρώνας δέχεται διάφορα σήματα, τότε υπολογίζει μία ποσότητα  $x$  από όλα τα δεδομένα που έχει και συγκρίνει την τιμή της ποσότητας αυτής με μια τιμή κατωφλίου,  $\theta$ , η οποία είναι χαρακτηριστική (σταθερή) και ορισμένη από την αρχή για τον νευρώνα αυτόν. Αν η τιμή της ποσότητας είναι μεγαλύτερη από την τιμή κατωφλίου, τότε λέμε ότι ο νευρώνας ενεργοποιείται. Αν όμως είναι μικρότερη, τότε ο νευρώνας παραμένει αδρανής, δηλ. στην δεδομένη στιγμή δεν μεταδίδει κανένα σήμα περαιτέρω στο δίκτυο. Επειδή ο νευρώνας εδώ δρα ως δυαδικό στοιχείο, γι' αυτό η έξοδος του,  $f(x)$ , θα είναι 1 όταν είναι ενεργοποιημένος και 0 όταν είναι αδρανής.

$$\begin{aligned} \text{Δηλαδή} \quad & f(x)=0 \text{ αν } x>\theta \\ \text{Και} \quad & f(x)=1 \text{ αν } x<\theta \end{aligned} \quad (3.1)$$

Με τον δεύτερο τρόπο δεν υπάρχει χαρακτηριστική τιμή κατωφλίου με την οποία γίνεται η σύγκριση της συνάρτησης (3.1). Η μετάδοση του σήματος γίνεται πάλι με την συνάρτηση  $f(x)$ , η οποία τώρα έχει μία ειδική μορφή. Χρησιμοποιούμε όλες τις τιμές των εισόδων και τις τιμές των βαρών,  $w$ , και υπολογίζουμε αριθμητικά την  $f(x)$ . Ένα παράδειγμα μορφής της συνάρτησης αυτής είναι το εξής:

$$f(x) = 1/1+e^{-x} \quad (3.2)$$

Η συνάρτηση (3.2) λέγεται σιγμοειδής συνάρτηση. Εκτός από τις μορφές (3.1) και (3.2) υπάρχουν και άλλες μορφές της  $f(x)$ . Η γενική της όμως ονομασία σε όλες τις περιπτώσεις είναι συνάρτηση μεταφοράς (transfer function), ή συνάρτηση ενεργοποίησης (activation function). Το κοινό χαρακτηριστικό που έχουν οι συναρτήσεις αυτές είναι ότι πρέπει να είναι πάντοτε μη- γραμμικές. Δεν αρκούν

γραμμικές συναρτήσεις, γιατί τότε η έξοδος θα ήταν ευθέως ανάλογη με την είσοδο, κάτι που δεν μπορεί να συμβεί στα νευρωνικά δίκτυα.

## Κεφάλαιο 4

### ΜΑΘΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

Ο πρωταρχικός σκοπός της λειτουργίας ενός τεχνητού νευρωνικού δικτύου είναι να μπορεί να λύνει συγκεκριμένα προβλήματα που του παρουσιάζουμε ή να επιτελεί από μόνο του ορισμένες διεργασίες, π.χ. να αναγνωρίζει εικόνες. Για να μπορεί όμως να γίνει αυτό λέμε ότι το νευρωνικό δίκτυο προηγουμένως πρέπει να εκπαιδευθεί. Αυτό είναι και το βασικό χαρακτηριστικό των νευρωνικών δικτύων. Τι ακριβώς όμως σημαίνει ότι ένα νευρωνικό δίκτυο εκπαιδεύεται;

Τα νευρωνικά δίκτυα δέχονται ορισμένες εισόδους και αντίστοιχα δίνουν ορισμένες εξόδους (input–output). Όταν λέμε εισόδους/εξόδους εννοούμε ότι παρουσιάζονται στο δίκτυο κάποια σήματα τα οποία έχουν αριθμητικές τιμές, π.χ. θα μπορούσε να είναι κάποιος δυαδικός αριθμός αποτελούμενος από 0 και 1. Οι αριθμοί αυτοί που δίνονται στην είσοδο του δικτύου αποτελούν κάποιο πρότυπο. Για ένα πρόβλημα μπορεί να απαιτούνται πολλά πρότυπα. Σε κάθε πρότυπο αντιστοιχεί και μία σωστή απάντηση, η οποία είναι το σήμα που πρέπει να πάρουμε στην έξοδο.

Η εκπαίδευση γίνεται με το να παρουσιάσουμε μια ομάδα από τέτοια 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 δηλαδή ποιος είναι ο στόχος, τι πρέπει να δίνει το δίκτυο ως απάντηση στα πρότυπα που του παρουσιάζουμε. Ουσιαστικά είναι σαν να δίνουμε στο δίκτυο μία ερώτηση και ακολούθως να του δίνουμε την απάντηση που αντιστοιχεί. Το δίκτυο χρησιμοποιεί την κατάλληλη συνάρτηση μεταφοράς  $f(x)$  για να μεταδίδει το σήμα σε όλη τη δομή του, από την είσοδο ως την έξοδο.

Κατά την διάρκεια της εκπαίδευσης το μόνο πράγμα που αλλάζει είναι οι τιμές των βαρών των συνδέσεων των νευρώνων. Αυτό δεν γίνεται πάντα με τον ίδιο τρόπο, αλλά εξαρτάται σημαντικά από την μέθοδο που χρησιμοποιούμε. Το δίκτυο με τα δεδομένα αυτά τροποποιεί την εσωτερική του δομή ώστε να μπορεί να κάνει την ίδια αντιστοιχία που του δώσαμε εμείς. Ενώ αρχικά ξεκινάει με τιμές στα βάρη  $w$  που είναι τυχαίες, κατά την διάρκεια της εκπαίδευσης μεταβάλλει τις τιμές αυτές, μέχρι να εκπαιδευθεί πλήρως. Ακολούθως, αφού βρει την σωστή εσωτερική δομή του, τότε θα μπορεί να λύνει και άλλα ανάλογα προβλήματα τα οποία δεν τα έχει δει



προηγούμενως, 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 τον τρόπο με τον οποίο τα δεδομένα παρουσιάζονται στο δίκτυο όταν αυτό εκπαιδεύεται. Γενικά, μπορούμε να πούμε ότι κατά την εκπαίδευση ενός δικτύου οι αλλαγές στα βάρη γίνονται με ένα από τους εξής δύο τρόπους:

- με εποπτευόμενο τρόπο
- με μη-εποπτευόμενο τρόπο(αυτο-εποπτευόμενο τρόπο)

Η εποπτευόμενη μάθηση είναι και ο πιο συχνός τρόπος στην εκπαίδευση των νευρωνικών δικτύων.

Αρχικά δίνουμε τις τιμές των εισόδων και των στόχων που πρέπει να μάθει το δίκτυο, δηλ. παρουσιάζουμε τα πρότυπα στο δίκτυο. Ξεκινούμε με τυχαίες τιμές στα βάρη  $w$ . Κατά την διαδικασία εκπαίδευσης το δίκτυο αλλάζει τις τιμές των βαρών διορθώνοντας αυτές ανάλογα με το σφάλμα που παίρνουμε (διαφορά από τον στόχο). Ο σκοπός μας εδώ είναι τελικά να ελαχιστοποιήσουμε την διαφορά (το σφάλμα) μεταξύ της επιθυμητής εξόδου και της τρέχουσας τιμής της εξόδου μετά από διαδοχικές αλλαγές των βαρών (ανακυκλώσεις διορθώσεων). Μερικές φορές η διαδικασία αυτή μπορεί να απαιτεί μεγάλους αριθμούς τέτοιων διορθώσεων και, επομένως, μεγάλους υπολογιστικούς χρόνους.

Στην μη-εποπτευόμενη εκπαίδευση απλώς δίνουμε την πληροφορία στο δίκτυο, αλλά δεν δίνουμε αντίστοιχους στόχους όπως προηγούμενως και έτσι δεν γίνεται κανένας έλεγχος ή σύγκριση για την πορεία του σφάλματος. Το δίκτυο δεν χρησιμοποιεί κάποια εξωτερική παράμετρο για την αλλαγή των βαρών. Υπάρχει συγκεκριμένη διαδικασία που ακολουθείται και καταλήγει σε εκπαίδευση του δικτύου. Το δίκτυο χρησιμοποιεί έναν εσωτερικό έλεγχο, ψάχνει να βρει κάποιες τάσεις ή κανονικότητα στα σήματα εισόδου και προσπαθεί έτσι ώστε οι έξοδοι να έχουν τα ίδια χαρακτηριστικά όπως και οι εισοδοί. Λέμε ότι έτσι έχουμε αυτό-εποπτευόμενη εκπαίδευση διότι το δίκτυο ελέγχει τον εαυτό του και διορθώνει τα σφάλματα στα δεδομένα με ένα μηχανισμό ανάδρασης (feedback). Ο τρόπος αυτός δεν συναντάτε τόσο συχνά όπως η εποπτευόμενη εκπαίδευση, αλλά είναι πολύ

χρήσιμος σε ορισμένες καταστάσεις που δεν υπάρχουν δεδομένα στο πρόβλημα. 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.

## Κεφάλαιο 5

### ΑΞΙΟΠΙΣΤΙΑ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ

Το Σχήμα 3.1 παρουσιάζει 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. Καθόσον, 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 κτλ. 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, μοιράζεται σε όλους τους επί μέρους νευρώνες. Έτσι λέμε 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.

Η αρχιτεκτονική των νευρωνικών δικτύων όμως διαφέρει από αυτήν των παραλλήλων επεξεργαστών, για το λόγο ότι οι απλοί επεξεργαστές των νευρωνικών δικτύων (δηλαδή οι νευρώνες) έχουν μεγάλο αριθμό διασυνδέσεων, ο οποίος συνολικά είναι πολύ μεγαλύτερος από τον αριθμό των νευρώνων. Και αυτό βέβαια γιατί κάθε νευρώνας έχει πολλές συνδέσεις. 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, δηλαδή ξέρουν μόνο 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. Οι πληροφορίες που αποθηκεύονται σε ένα νευρωνικό δίκτυο μοιράζονται σε ένα μεγάλο αριθμό μονάδων, δηλαδή 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, παρόλο 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 η τεχνική αυτή μπορεί να χρησιμοποιείται 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 οι συνηθισμένες τιμές σε διάφορα προβλήματα που παρουσιάζονται για την μεγαλύτερη δυνατή ανοχή 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 συνηθεις 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 για αυτό που θέλουμε. 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.

Στον παρακάτω πίνακα υπάρχουν 8 ιδιότητες που περιγράφουν με μικρές προτάσεις τις ομοιότητες και διαφορές μεταξύ των γνωστών μας υπολογιστών και των νευρωνικών δικτύων.

**Πίνακας 5.1:** Διαφορές ανάμεσα στα τεχνητά νευρωνικά δίκτυα και στους Η/Υ

Νευρωνικά δίκτυα		Υπολογιστές
1	Εργάζονται με σύγχρονο τρόπο λειτουργίας	Εργάζονται με ασύγχρονο τρόπο λειτουργίας
2	Παράλληλη επεξεργασία	Σειριακή επεξεργασία
3	Εκπαίδευση με παραδείγματα αλλάζοντας τα βάρη των συνδέσεων τους	Προγραμματίζεται με εντολές λογικού χαρακτήρα (if - then)
4	Η μνήμη, τα δίκτυα και οι μονάδες λειτουργίας συνυπάρχουν	Η μνήμη και επεξεργασία πληροφορίας χωρίζονται
5	Ανοχή στα σφάλματα	Καμιά ανοχή στα σφάλματα
6	Αυτό-οργάνωση κατά την διαδικασία της εκπαίδευσης	Εξαρτάται εξ ολοκλήρου από το προσφερόμενο λογισμικό
7	Η πληροφορία αποθηκεύεται στα βάρη των συνδέσεων	Η πληροφορία αποθηκεύεται σε συγκεκριμένες διευθύνσεις μνήμης
8	Ο χρόνος ενός κύκλου είναι της τάξης msec	Ο χρόνος ενός κύκλου είναι της τάξης nsec

Στην συνέχεια αναφέρονται βασικά πλεονεκτήματα και μειονεκτήματα των τεχνητών νευρωνικών δικτύων (ΤΝΔ).

Τα πλεονεκτήματα από τη χρήση των Τεχνητών Νευρωνικών Δικτύων στα διάφορα προβλήματα, είναι τα εξής:

- *Μη – γραμμικότητα.*

Ένα Τεχνητό Νευρωνικό Δίκτυο δομείται από τη σύνδεση νευρώνων, οι οποίοι είναι μη γραμμικές συσκευές. Η μη-γραμμικότητα είναι πολύ σημαντική ιδιότητα, αν ο φυσικός μηχανισμός για την παραγωγή των σημάτων εισόδου είναι μη-γραμμικός.

- *Σχεδιασμός Εισόδου – Εξόδου.*

Το δίκτυο μαθαίνει με τον κατάλληλο σχεδιασμό εισόδου-εξόδου. Γιατί στην επιβλεπόμενη μάθηση, εφαρμόζεται ένα σύνολο δειγμάτων εξάσκησης και κάθε παράδειγμα αποτελείται από ένα σήμα εισόδου και την επιθυμητή απόκριση.

- *Προσαρμοστικότητα.*

Τα Τεχνητά Νευρωνικά Δίκτυα έχουν τη δυνατότητα να προσαρμόζουν τους συντελεστές βαρύτητας τους στις αλλαγές του περιβάλλοντός τους. Δηλαδή, με αλλαγή αρχιτεκτονικής ή δεδομένων εισόδου υπάρχει αυτόματη αλλαγή και των συντελεστών βαρύτητας του Τεχνητού Νευρωνικού Δικτύου.

- *Αποδεκτή Απόκριση.*

Ένα Τεχνητό Νευρωνικό Δίκτυο σχεδιάζεται για να παρέχει πληροφορίες όχι μόνο για το συγκεκριμένο υπόδειγμα που επιλέγεται αλλά και για την εμπιστοσύνη στην απόφαση που παίρνεται στο εκάστοτε πρόβλημα.

- *Συναφής Πληροφορία*

Η γνώση αναπαριστάνεται από την πολύ δομημένη και ενεργή κατάσταση του Τεχνητού Νευρωνικού Δικτύου.

- *Ομοιομορφία Ανάλυσης και Σχεδιασμού.*

Η έννοια είναι ότι ο ίδιος συμβολισμός χρησιμοποιείται σε όλα τα πεδία που περιέχουν εφαρμογή των νευρωνικών δικτύων.

Τα μειονεκτήματα από τη χρήση των Τεχνητών Νευρωνικών Δικτύων, παρουσιάζονται παρακάτω:

- Το βασικό μειονέκτημα των νευρωνικών δικτύων είναι ότι συμπεριφέροντε καθαρά ως «μαύρο κουτί». Δηλαδή, δεν παρέχουν τη μαθηματική εξίσωση λύσης, η οποία πολλές φορές ίσως είναι προτιμότερη.

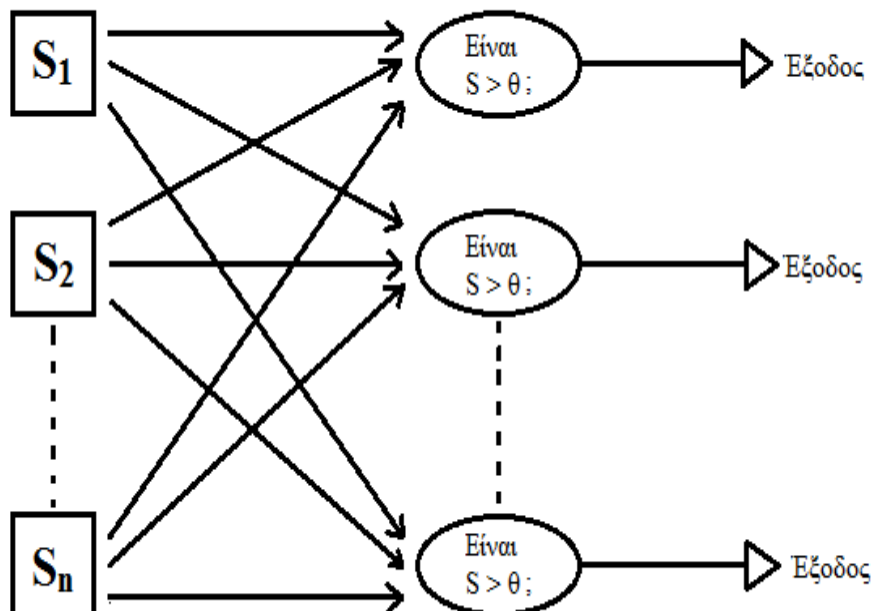
- Υπάρχει η πιθανότητα πολύ υψηλής απόδοσης για τα δεδομένα που έχει εκπαιδευτεί, αλλά να μην αποδίδει το ίδιο καλά σε νέα δεδομένα (φαινόμενο υπέρ-μνημόνευσης - over fitting).

- Δεν έχει βρεθεί ένας ενιαίος βέλτιστος τρόπος για την αρχιτεκτονική των τεχνητών νευρωνικών δικτύων. Δηλαδή, για το ποιός είναι ο καταλληλότερος αριθμός νευρώνων σε κάθε επίπεδο ή ο καταλληλότερος αριθμός των κρυφών επιπέδων κλπ. Επομένως, για κάθε πρόβλημα πραγματοποιούνται πολλές δοκιμές μέχρις ότου να βρεθεί το βέλτιστο και πιο αξιόπιστο Τεχνητό Νευρωνικό Δίκτυο. Άρα, η συγκεκριμένη μεθοδολογία απαιτεί χρόνο και εμπειρία από τον χρήστη.

## Κεφάλαιο 6

### ΠΡΟΤΥΠΑ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

Στο δεύτερο κεφάλαιο είδαμε την απλή δομή ενός νευρωνικού δικτυού και εξηγήσαμε την συνδεσμολογία του και τις ιδιότητες του. Σε αυτό το κεφάλαιο θα αναφέρουμε μερικά βασικά πρότυπα αισθητήρων νευρωνικών δικτύων και θα τα αναλύσουμε. Το μοντέλο του αισθητήρα (perceptron) είναι από τα πρώτα μοντέλα νευρωνικών δικτύων που αναπτύχθηκαν και έδωσαν στην περιοχή αυτή μεγάλη ώθηση χάρη στις επιτυχίες που είχε από την αρχή. Πολλά δίκτυα που αναπτύχθηκαν αργότερα, πιο περίπλοκα, ξεκίνησαν από την βάση του αισθητήρα. Σήμερα, υπάρχουν πολλές παραλλαγές νευρωνικών δικτύων που βασίζονται στον αισθητήρα και έχουν διαφορετικές δομές, άλλες απλές και άλλες περίπλοκες. Η πιο απλή μορφή είναι ο λεγόμενος στοιχειώδης αισθητήρας (elementary perceptron), Σχήμα 3.1, γιατί αποτελείται από ένα μόνο νευρώνα και είναι το πιο απλό, αυτοδύναμο σύστημα που υπάρχει και επιτελεί μία ορισμένη διεργασία.



Σχημα 6.1: Αισθητήρας με  $n$  νευρώνες

Ανεβαίνοντας ως προς την πολυπλοκότητα, έχουμε νευρωνικά δίκτυα τα οποία έχουν πολλούς νευρώνες, όπως στο Σχήμα 6.1, οι οποίοι είναι οργανωμένοι σε δύο επίπεδα, ένα επίπεδο στο οποίο εισέρχονται τα σήματα (επίπεδο εισόδου) και ένα επίπεδο όπου βγαίνει το αποτέλεσμα του νευρωνικού δικτύου (επίπεδο εξόδου). Στον στοιχειώδη αισθητήρα, ο μοναδικός νευρώνας του συστήματος έχει έναν ορισμένο αριθμό συνδέσεων που προέρχονται από άλλους νευρώνες (τους οποίους όμως δεν εξετάζουμε), όπως φαίνονται στο Σχήμα 3.1.

Ο νευρώνας είναι ο κύκλος και οι συνδέσεις του είναι οι διάφορες γραμμές. Έχει ένα ορισμένο αριθμό εισόδων αλλά μία μόνο έξοδο. Αυτό σημαίνει ότι η μονάδα αυτή δέχεται πολλές εισόδους,  $s_1, s_2, s_3$  κτλ. Αλλά παράγει μία μόνο έξοδο, που όπως φαίνεται στο Σχήμα είναι στα δεξιά του κύκλου. Στο σημείο αυτό χρειάζεται προσοχή: μπορεί ένας νευρώνας να έχει πολλές εξόδους και άρα πολλές γραμμές δεξιά του κύκλου όλες όμως οι εξοδοί αυτές πάντα έχουν την ίδια τιμή. Αν λοιπόν, υπάρχουν πολλές γραμμές- έξοδοι, ποτέ δεν μπορούν να έχουν διαφορετικές τιμές. Αυτό λοιπόν εννοούμε όταν λέμε ότι ένας νευρώνας έχει μία μόνο έξοδο.

Κάθε εισερχόμενο σήμα,  $s_i$ , συνδέεται με τον κεντρικό νευρώνα με ένα βάρος  $w_i$ . Το βάρος  $w$  μας δείχνει κατά κάποιο τρόπο την αλληλεπίδραση μεταξύ των δύο νευρώνων τους οποίους συνδέει. Στην απλή περίπτωση του ενός νευρώνα, μπορούμε να πούμε ότι το  $w$  είναι η επίδραση του εισερχόμενου σήματος με τον νευρώνα αυτό. Αυτό που έχει σημασία δεν είναι η τιμή του βάρους  $w$  από μόνη της ούτε η τιμή του σήματος  $s$ , αλλά είναι το γινόμενο  $s_i w_i$ . Έτσι κάθε  $s_i$  πολλαπλασιάζεται επί το βάρος  $w_i$  που έχει η σύνδεση  $i$  και τελικά αυτό που παρουσιάζεται στον νευρώνα από κάθε εισερχόμενο σήμα είναι το γινόμενο  $s_i w_i$ . Ο αισθητήρας κατόπιν αθροίζει τα γινόμενα αυτά για όλους τους  $n$  όρους (όπου  $n$  είναι ο αριθμός των εισόδων) και θεωρούμε, λοιπόν, ότι λαμβάνει ένα συνολικό σήμα με τιμή:

$$S = \sum_i^n s_i \cdot w_i$$

Μερικές φορές θεωρούμε ότι εκτός από τα εισερχόμενα σήματα και τα αντίστοιχα βάρη  $w$ , ο νευρώνας έχει και ένα εσωτερικό βάρος που τον χαρακτηρίζει, το οποίο πρέπει να ληφθεί υπόψη στην εξίσωση 6.1. Το εσωτερικό αυτό βάρος λέγεται bias,  $b$ , ή αλλιώς προδιάθεση ή παράγων προδιάθεσης του νευρώνα. Το βάρος αυτό είναι τελείως ξεχωριστό από τα άλλα βάρη, αλλά δρα με τον ίδιο τρόπο όπως τα άλλα βάρη  $w$  που είδαμε μέχρι τώρα. Θεωρούμε πάντοτε ότι η τιμή του σήματος που



περνάει σε όλα τα εσωτερικά βάρη είναι 1. Έτσι, λοιπόν, οι μονάδες του  $b$  θα είναι οι ίδιες με τις μονάδες του γινομένου ( $s \cdot w$ ), ώστε η εξίσωση 6.1 στην πιο γενική της μορφή γίνεται τώρα:

$$S = b + \sum_i^n \cdot s_i \cdot w_i \quad (6.2)$$

Ο όρος  $b$  δεν έχει καμία φυσική σημασία και δεν πρέπει να αποδίδεται στο εσωτερικό του νευρώνα στον οποίο καταλήγει. Μερικές φορές θεωρείται ότι είναι ένα εξωτερικό ερέθισμα, το οποίο προστίθεται στο υπόλοιπο άθροισμα για να δώσει το σωστό  $S$ . Όλα τα προβλήματα δεν αντιμετωπίζονται με τον ίδιο τρόπο και έτσι άλλες φορές χρησιμοποιείται η εξίσωση 6.1 (χωρίς τον όρο  $b$ ) και άλλες η εξίσωση 6.2 (με τον όρο  $b$ ), χωρίς να υπάρχει ένας γενικός κανόνας. Ακολουθώντας εφαρμόζουμε την συνάρτηση κατωφλίου Heaviside, με μία συγκεκριμένη τιμή του κατωφλίου,  $\theta$ . Αν το σήμα είναι ίσο ή ξεπερνά το κατώφλι δυναμικού, τότε ο νευρώνας διεγείρεται και είναι έτοιμος να στείλει έναν παλμό που έχει πάντοτε το ίδιο μέγεθος. Αμέσως μετά τον παλμό, ο νευρώνας επανέρχεται στην αρχική του κατάσταση. Αν οι συνθήκες το επιτρέψουν, μπορεί αργότερα να ενεργοποιηθεί πάλι. Το σήμα όμως που μεταδόθηκε συνεχίζει την ίδια διαδικασία σε άλλους νευρώνες του δικτύου χωρίς να ελαττωθεί καθόλου. Μεταδίδεται προς μία κατεύθυνση, δηλαδή απομακρύνεται από το σώμα του κυττάρου.

Η συνάρτηση Heaviside είναι ίδια με την εξίσωση 3.1 που είδαμε στο προηγούμενο κεφάλαιο, και αυτό διότι η  $f(x)$ , η συνάρτηση μεταφοράς, είχε την ίδια περιγραφή όπως και εδώ. Συγκρίνουμε το  $\theta$  με το άθροισμα  $S$ . Εάν  $S > \theta$ , τότε ο αισθητήρας ενεργοποιείται και θεωρούμε ότι πυροδοτεί. Εάν  $S < \theta$ , τότε το άθροισμα  $S$  μηδενίζεται και ο αισθητήρας παραμένει αδρανής. Αυτό συνοψίζεται ως:

$$\text{Εάν } S > \theta \text{ τότε η τιμή της εξόδου} = 1 \quad (6.3)$$

$$\text{Εάν } S < \theta \text{ τότε η τιμή της εξόδου} = 0 \quad (6.4)$$

Αντιλαμβανόμαστε λοιπόν ότι η ενεργητικότητα του αισθητήρα εξαρτάται από τρεις παραμέτρους: τα βάρη των συνδέσεων, τις τιμές των εισόδων και την τιμή του κατωφλίου. Θεωρούμε ότι αυτό που μαθαίνει το σύστημά μας αποθηκεύεται στα βάρη των συνδέσεων, τα οποία, όπως θα δούμε και παρακάτω, μεταβάλλονται συνεχώς κατά την διάρκεια που το σύστημα μαθαίνει κάποια πληροφορία. Ξεκινώντας από το πρότυπο του στοιχειώδους αισθητήρα μπορούμε να αναπτύξουμε προχωρημένα πρότυπα που περιέχουν αναγκαστικά περισσότερους του ενός νευρώνες.

Ένας αισθητήρας με περίπλοκη δομή δίνεται στο Σχήμα 6.1. Στο Σχήμα αυτό οι κόμβοι εισόδου παρίστανται με ένα τετράγωνο. Στους κόμβους αυτούς δεν γίνεται καμία επεξεργασία του σήματος, αλλά χρησιμεύουν μόνο για να δέχονται το σήμα. Οι υπόλοιποι κόμβοι τους οποίους γίνεται επεξεργασία του σήματος δηλώνονται με ένα κύκλο. Εδώ έχουμε  $n$  νευρώνες, αντί για έναν μόνο που έχει ο στοιχειώδης αισθητήρας. Στην πιο γενική περίπτωση έχουμε πλήρη συνδεσμολογία, δηλαδή κάθε εισερχόμενο σήμα  $s_i$  παρουσιάζεται και στους  $n$  νευρώνες, με διαφορετικό βάρος κάθε φορά. Η διαδικασία σύγκρισης με το κατώφλι  $\theta$  είναι η ίδια όπως και στο απλό μοντέλο, αλλά εδώ έχουμε μια πλειάδα από εξόδους των οποίων ο αριθμός είναι  $n$ , όσο δηλ. και ο αριθμός των νευρώνων.

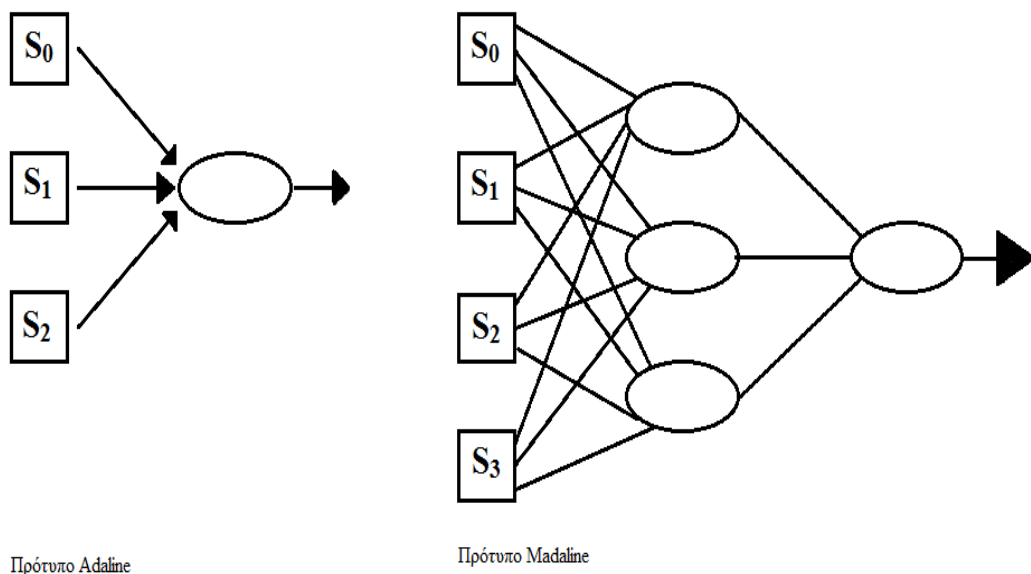
Υπάρχουν και διάφορα άλλα παρόμοια μοντέλα, που ονομάζονται επίσης συλλογικά αισθητήρες, και μερικά είναι πιο περίπλοκα από τα παραπάνω, αλλά ο μηχανισμός λειτουργίας τους είναι ο ίδιος όπως αυτός που είδαμε. Δυο από τα σημαντικότερα πρότυπα αισθητήρων νευρωνικών δικτύων είναι τα πρότυπα Adaline και Madaline.

Το adaline βγαίνει από το adaptive linear neuron (ada-line), ενώ το madaline από το multilayer adaline (m-adaline). Παρουσιάστηκαν από τον B. Widrow το 1959. Στο πρώτο πρότυπο, το adaline, έχουμε ένα δίκτυο με πολλές εισόδους και μία έξοδο, (Σχήμα 6,2). Κάθε είσοδος έχει το δικό της βάρος. Υπάρχει επίσης ο στόχος και έτσι η έξοδος συγκρίνεται κάθε φορά με τον στόχο, ώστε να βρεθεί η τιμή του σφάλματος. Οι νευρώνες εδώ έχουν σήμα με τιμές  $+1$  και  $-1$ . Επίσης η έξοδος του δικτύου πρέπει να είναι  $+1$  ή  $-1$ , κατ' αντίθεση με τον απλό αισθητήρα, όπου οι δυαδικές τιμές ήταν  $0$  ή  $1$ . Παράγεται το άθροισμα της εξίσωσης 3.1. Το πρώτο σήμα που εισέρχεται στην είσοδο έχει πάντα (εξ ορισμού) την τιμή  $+1$ , δηλαδή  $s_0 = 1$ . Το αντίστοιχο βάρος  $w$  μεταβάλλεται όπως και τα άλλα βάρη. Μετά την άθροιση γίνεται η σύγκριση με το  $\theta$ , και η έξοδος είναι:

$$\text{Εάν } S \geq \theta \text{ τότε έξοδος} = 1 \quad (6,5)$$

$$\text{Εάν } S < \theta \text{ τότε έξοδος} = -1 \quad (6,6)$$

Δηλαδή εδώ έχουμε  $\theta = 0$ .



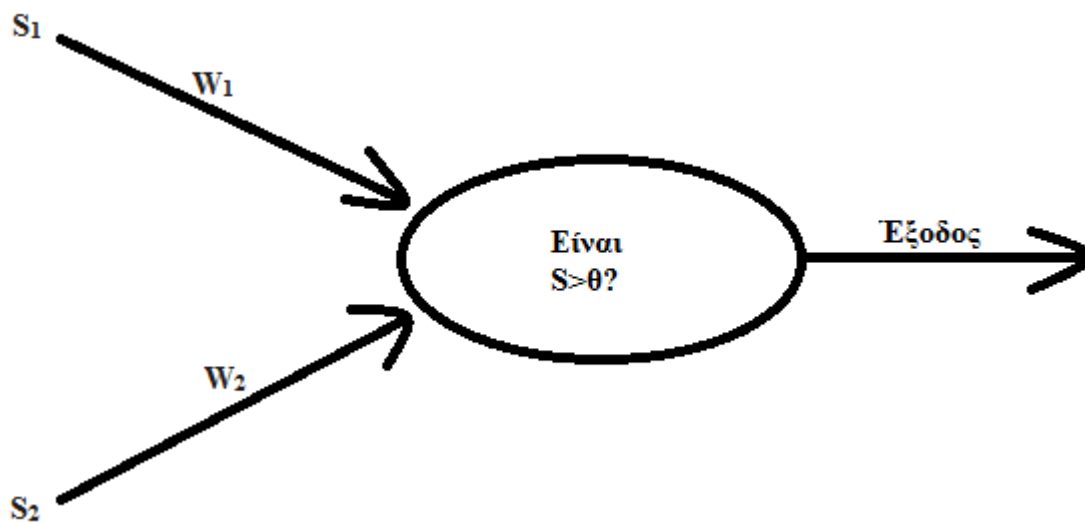
**Σχήμα 6.2:** Πρότυπα νευρωνικών δικτύων adaline και madaline

Στο πρότυπο madaline έχουμε ένα επίπεδο μονάδων adaline (στο Σχήμα 6.2 οι μονάδες αυτές είναι 3) που ενώνονται με μία μονάδα εξόδου, σχηματίζοντας έτσι μία μονάδα madaline. Εφόσον έχουμε εδώ 3 επίπεδα, οι μεταβολές στα  $w$  δεν γίνονται σε όλα, αλλά μόνο σε αυτά από την είσοδο (4 νευρώνες) στο μεσαίο επίπεδο (3 νευρώνες). Χρησιμοποιούμε, δηλαδή, μία διαφορετική συνάρτηση μεταφοράς εδώ. Από το μεσαίο επίπεδο προς την έξοδο η απόφαση λαμβάνεται πλειοψηφικά ή μπορεί να χρησιμοποιηθεί και άλλη λογική συνάρτηση. Αν περισσότεροι από τους μισούς νευρώνες δίνουν  $+1$  ή  $-1$ , τότε η έξοδος madaline είναι επίσης  $+1$  ή  $-1$ . Μπορεί όμως να χρησιμοποιηθεί και άλλη λογική συνάρτηση για την λήψη της απόφασης. Η διαδικασία εκπαίδευσης έχει επίσης εδώ την ίδια μορφή. Η έξοδος συγκρίνεται με τον στόχο. Από την σύγκριση προκύπτει ένα σφάλμα, το οποίο χρησιμοποιείται για την μεταβολή των  $w$ . Σε μία δεδομένη στιγμή μόνο μία μονάδα adaline μεταβάλλει τα βάρη της.

Ένα από τα πιο γνωστά προβλήματα που επιλύονται με νευρωνικά δίκτυα είναι αυτό της εκμάθησης της συνάρτησης X-OR (exclusive-or), δηλ. της συνάρτησης της αποκλειστικής διάζευξης, όπως λέγεται. Η συνάρτηση αυτή δέχεται δύο εισόδους και δίνει μία έξοδο. Οι εισοδοί και η έξοδος μπορεί να είναι 0 ή 1 μόνον και ισχύει ο εξής περιορισμός: εάν και οι δύο εισοδοί είναι ίδιες, τότε η έξοδος είναι 0, εάν είναι διαφορετικές, τότε η έξοδος είναι 1. Οι όροι αυτοί συνοψίζονται στον Πίνακα 3.1, που ονομάζεται και πίνακας αληθείας της συνάρτησης.

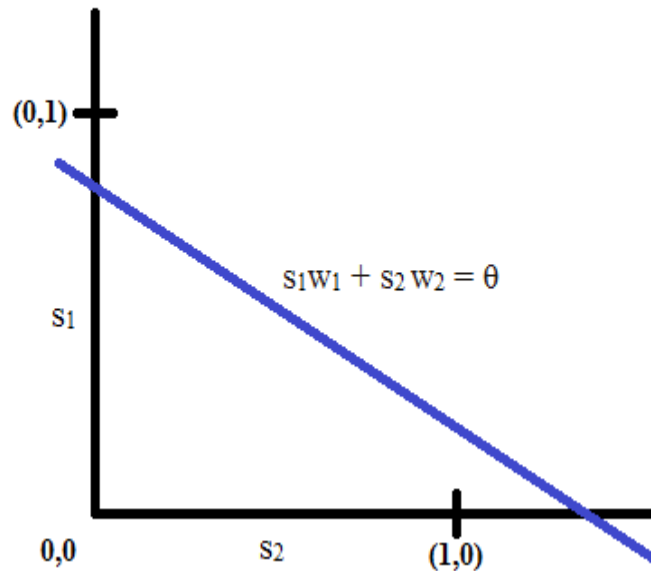
**Πίνακας 6.1:** Η συνάρτηση X-OR

Είσοδος 1	Είσοδος 2	Έξοδος
0	0	0
0	1	1
1	0	1
1	1	0



**Σχήμα 6.3:** Ο αισθητήρας με δυο εισόδους για το πρόβλημα X-OR

Εάν χρησιμοποιήσουμε τον στοιχειώδη αισθητήρα με δύο εισόδους και μία έξοδο τότε έχουμε το Σχήμα 6.3 που είναι μία ειδική περίπτωση του Σχήματος 3.1. Στο δίκτυο του Σχήματος 6.3 κάθε φορά που έρχονται οι δύο εισόδοι  $s_1$  και  $s_2$  τίθεται το ερώτημα της σύγκρισης μεταξύ του  $S$  και του  $\theta$ . Θέλουμε το δίκτυο να δίνει έξοδο  $= 0$ , όταν το  $S$  είναι  $S < 0,5$  και να δίνει έξοδο  $= 1$ , όταν  $S > 0,5$ . Βλέπουμε όμως καθαρά ότι δεν υπάρχει κανένας συνδυασμός τιμών των  $w_1$  και  $w_2$  που να παράγει τις σχέσεις που περιλαμβάνονται στον Πίνακα 6.1. Ας θεωρήσουμε ότι το κατώφλι  $\theta = 0,5$ . Η αλγεβρική σχέση γίνεται  $s_1w_1 + s_2w_2 = 0,5$  και περιγράφει το δίκτυό μας. Η εξίσωση αυτή είναι γραμμική ως προς  $s_1$  και  $s_2$ . Αυτό σημαίνει ότι όλες οι τιμές των  $s_1$  και  $s_2$  που ικανοποιούν την εξίσωση αυτή θα βρίσκονται σε μία ευθεία γραμμική στο επίπεδο  $x-y$ , όπως π.χ. η ευθεία του Σχήματος 6.4.



**Σχήμα 6.4:** Το πρόβλημα της συνάρτησης X-OR σε αναπαράσταση στο επίπεδο x-y

Οποιαδήποτε τιμή αν έχουν τα  $s_1$  και  $s_2$  πάνω στην γραμμή αυτή, θα δώσουν  $S = 0,5$ . Εάν τα  $s_1$  και  $s_2$  βρίσκονται στην μία πλευρά της γραμμής, τότε το  $S > \theta$ , και έξοδος = 1. Αν τα  $s_1$  και  $s_2$  βρίσκονται στην άλλη πλευρά της γραμμής, τότε  $S < \theta$  και έξοδος = 0. Αλλάζοντας τις τιμές  $w_1$  και  $w_2$  καθώς και την τιμή του  $\theta$ , θα αλλάξει η κλίση και η θέση της γραμμής αυτής. Αυτό που θέλουμε εμείς όμως είναι τα σημεία  $(0,0)$  και  $(1,1)$  να βρίσκονται από τη μία πλευρά της ευθείας και τα σημεία  $(0,1)$  και  $(1,0)$  να βρίσκονται από την άλλη. Μόνον τότε το δίκτυο θα δίνει την σωστή απάντηση. Βλέπουμε ότι δεν υπάρχει κανένας τρόπος να τραβήξουμε μία ευθεία γραμμή, με οποιαδήποτε κλίση, που να ικανοποιεί την συνθήκη αυτή. Καταλήγουμε λοιπόν στο συμπέρασμα ότι το δίκτυο του Σχήματος 6.3 με ένα μόνο κόμβο, ανεξάρτητα από τιμές  $w_1, w_2$ , και  $\theta$ , δεν μπορεί να λύσει το πρόβλημα της συνάρτησης X-OR και χρειάζεται νευρωνικό δίκτυο μεγαλύτερης πολυπλοκότητας.

## Κεφάλαιο 7

### ΕΦΑΡΜΟΓΕΣ

Οι εφαρμογές των τεχνητών νευρωνικών δικτύων σε διάφορους τομείς της Επιστήμης και της Τεχνολογίας είναι πολλές. Οι περισσότερες από αυτές τις εφαρμογές έχουν ήδη υλοποιηθεί και αποτελούν εμπορικά προϊόντα.

- **Αεροπορία:** Δημιουργία αυτόματων πιλότων και προγραμμάτων προσομοίωσης πτήσης, συστήματα ελέγχου πτήσης, ανίχνευση ελαττωμάτων σε τμήματα των αεροπλάνων.
- **Άμυνα:** Χειρισμός μη επανδρωμένων οχημάτων και αεροπλάνων, αναγνώριση σημάτων από radar, δημιουργία «έξυπνων όπλων», αναγνώριση και σκόπευση στόχων, βελτιστοποίηση αξιοποίησης αποθεμάτων, κρυπτογραφία, πλοήγηση όπλων, νέα είδη αισθητήρων, σόναρ, ψηφιακή επεξεργασία σημάτων, συμπίεση δεδομένων, εξαγωγή χαρακτηριστικών, αναγνώριση σήματος / εικόνας.
- **Αυτοκίνηση:** Αυτοκινούμενα συστήματα αυτόματης πλοήγησης.
- **Βιολογία:** Βοήθεια στην κατανόηση του εγκεφάλου και άλλων συστημάτων, δημιουργία μοντέλων αμφιβληστροειδούς χιτώνα και κοχλία.
- **Βιομηχανία:** Βιομηχανικός έλεγχος διεργασιών, ανάλυση και σχεδίαση προϊόντων, συστήματα ποιοτικού ελέγχου, διάγνωση βλαβών διεργασιών και μηχανών, ανάλυση σχεδιασμού χημικών προϊόντων, δυναμική μοντελοποίηση συστημάτων χημικών διεργασιών.
- **Γεωλογία:** Ανάλυση πιθανότητας ύπαρξης πετρελαίου σε γεωλογικούς σχηματισμούς, εντοπισμός φυσικού αερίου. ανάλυση πετρωμάτων σε ορυχεία, ανάλυση της μόλυνσης του περιβάλλοντος.
- **Επιχειρήσεις:** Αξιολόγηση υποψηφίων για κάποια θέση, βελτιστοποίηση του συστήματος κράτησης θέσεων σε μεταφορικά μέσα, αναγνώριση γραφικού χαρακτήρα.
- **Ηλεκτρονική:** Πρόβλεψη ακολουθίας κωδίκων, μορφοποίηση ολοκληρωμένων κυκλωμάτων, έλεγχος διεργασιών, διάγνωση βλαβών ολοκληρωμένων κυκλωμάτων, μηχανική όραση.
- **ΙΑτρική:** Ανάλυση ομιλίας για την κατασκευή ακουστικών βοηθημάτων, διάγνωση βασισμένη στα συμπτώματα, έλεγχος χειρουργείου, εξαγωγή

συμπερασμάτων από ακτινογραφίες, ανάλυση καρδιογραφημάτων και εγκεφαλογραφημάτων, εντοπισμός καρκίνου σε μαστογραφίες.

- **Κατασκευές:** Αυτόματος έλεγχος, έλεγχος γραμμής παραγωγής, έλεγχος ποιότητας, επιλογή τμημάτων κατά το στάδιο της συναρμολόγησης.
- **Μεταφορές:** Συστήματα διάγνωσης βλαβών φρένων, χρονοπρογραμματισμός οχημάτων, συστήματα δρομολόγησης.
- **Περιβάλλον:** Πρόγνωση του καιρού, ανάλυση τάσεων και καιρικών συνθηκών.
- **Ρομποτική:** Έλεγχος τροχιάς και σύστημα όρασης ρομπότ.
- **Τηλεπικοινωνίες:** Συμπύση εικόνας και δεδομένων, αυτοματοποιημένες υπηρεσίες πληροφοριών, μετάφραση πραγματικού χρόνου, συστήματα επεξεργασίας πληρωμών.
- **Τραπεζικές εφαρμογές:** Υπολογισμός κινδύνου σε δάνεια και υποθήκες, αναγνώριση πλαστογραφιών, μετάφραση χειρόγραφων εντύπων, εκτίμηση τιμών μετοχών και συναλλάγματος.
- **Υπολογιστές:** Αναγνώριση ομιλίας, εντοπισμός φωνηέντων φθόγγων, μετατροπή κειμένου σε ομιλία, δρομολόγηση πληροφοριών σε δίκτυα υπολογιστών.

Εκτός από τα παραπάνω, την τελευταία δεκαετία, τα νευρωνικά δίκτυα έχουν αρχίσει να βρίσκουν εφαρμογή σε προβλήματα Υπολογιστικής Στατικής και Δυναμικής των κατασκευών. Τέτοιου είδους εφαρμογές είναι η ανάλυση αξιοπιστίας κατασκευών όπου στόχος είναι η πρόβλεψη των αποτελεσμάτων της ανάλυσης, ή προβλήματα βέλτιστου σχεδιασμού όπου οι απαιτούμενες τιμές των περιορισμών για κάθε νέο σχεδιασμό προέρχονται από την πρόβλεψη ενός κατάλληλα εκπαιδευμένου νευρωνικού δικτύου. Επίσης εφαρμόζονται σε προβλήματα θραυστομηχανικής, καθώς επίσης και σε προβλήματα προσαρμοστικών και στοχαστικών πεπερασμένων στοιχείων. Ένα σωστά εκπαιδευμένο νευρωνικό δίκτυο μπορεί να παράγει αποδεκτά, από πλευράς ακρίβειας, αποτελέσματα σε σύντομο υπολογιστικό χρόνο. Η ιδιότητα αυτή των νευρωνικών δικτύων αποτελεί και το βασικό τους πλεονέκτημα. Η προσέγγιση της λύσης μέσω ενός νευρωνικού δικτύου έχει μεγάλη αξία σε περιπτώσεις χρονοβόρων αναλύσεων όπου είναι αναγκαία μια γρήγορη εκτίμηση της συμπεριφοράς των φορέων. Από το παραπάνω πλήθος εφαρμογών, μπορούμε να συμπεράνουμε ότι τα νευρωνικά δίκτυα έχουν αναπτυχθεί κατά κόρον.

## Κεφάλαιο 8

### ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΝΑΥΤΙΛΙΑ

Τα τεχνητά νευρωνικά δίκτυα εκτός από τούς τομείς που αναφερθήκαν στο προηγούμενο κεφάλαιο, έχουν εξελιχθεί αρκετά και σε διάφορες ειδικότητες (τεχνικές, οικονομικές κλπ.) του τομέα της ναυτιλίας όπου θα τα εξετάσουμε παρακάτω. Όσον αφορά στην πρόβλεψη οικονομικών μεγεθών, αρκετοί έχουν επιχειρήσει να σχεδιάσουν ένα συμβουλευτικό σύστημα που θα έπαιρνε αποφάσεις για κερδοφόρες κινήσεις στην αγορά.

#### **8.1 Ναυτιλιακά Παράγωγα.**

Τα Νευρωνικά Δίκτυα χρησιμοποιούνται στην πρόγνωση της απόδοσης και εξέλιξης των Ναυτιλιακών Παραγώγων ή Freight Forward Agreements (FFAs) και γενικά στην πορεία της Ναυτιλιακής αγοράς και μπορούν να αποτελέσουν ανεκτίμητο εργαλείο, ικανό να συνεισφέρει αποφασιστικά στην αποτελεσματικότητα των επενδύσεων. Από την άλλη πλευρά, η πρόβλεψη πρέπει να έχει πάντα βοηθητικό ρόλο, να χρησιμοποιείται με σύνεση και να αποφεύγεται η αλόγιστη χρήση των αποτελεσμάτων αυτής. Δεν πρέπει άλλωστε να ξεχνάμε, ότι τα μοντέλα αυτά διδάσκονται από το παρελθόν ώστε να προβλέπουν το μέλλον, το οποίο μπορεί να επιφυλλάσει εκπλήξεις. Πρέπει να αποφεύγεται η κατάχρησή τους. Ιδανικώς, συνιστάται η χρήση τους να γίνεται σε συνδυασμό με την εμπειρική γνώση και όχι αλόγιστα, για την αποφυγή αρνητικών αποτελεσμάτων.

Σημαντικότερο προσόν που επιδεικνύουν τα Τεχνητά Νευρωνικά Δίκτυα είναι ότι μπορούν συνεχώς εμπλουτίζουν τη γνώση τους με νέες πληροφορίες που προκύπτουν από νέες συνθήκες στην αγορά. Εύκολα δηλαδή μπορούν να αφομοιώνουν καινούργια δεδομένα, ανακαλύπτοντας καινούργιους δεσμούς αλληλεπίδρασης αυτών με τα επιθυμητά δεδομένα και τελικά να παράγουν αναβαθμισμένα αποτελέσματα. Πάλι όμως εδώ, εισάγεται ως καθοριστικός παράγων η έννοια της μεταβλητότητας της μελετώμενης αγοράς, αφού όσο μεγαλύτερη είναι αυτή, για τόσο μικρότερο χρονικό διάστημα θα έχουν αξία τα νέα δεδομένα, έτσι ώστε να γίνει κατανοητή η χρήση ενός νευρωνικού δικτύου στην πρόγνωση της απόδοσης των ναυτιλιακών παραγώγων.



## **8.2 Εργασίες χειρισμού φορτίων σε λιμάνια**

Εκτός από το έργο που προσφέρουν στον οικονομικό τομέα, τα νευρωνικά δίκτυα χρησιμοποιούνται σε διάφορες εργασίες χειρισμού των φορτίων σε μεγάλα λιμάνια όπως του Rotterdam. Ο χειρισμός είναι μια διαδικασία που προκαλεί αίσθηση. Είναι μια πλήρης αυτοματοποιημένη και εκπαιδευμένη διαδικασία που λειτουργεί σύμφωνα με τις τελευταίες εξελίξεις στην ρομποτική και στα νευρωνικά δίκτυα σε συνδυασμό με εφαρμογές αισθητηρίων, οι οποίες λειτουργούν σύμφωνα με τα τελευταία δεδομένα ασφαλείας. Οι γερανοί ως επί το πλείστον είναι βραχίονες που στο τέλος έχουν αρπαγές οι οποίες χρησιμοποιούν εφαρμογές οπτικών νευρώνων. Η διαδικασία στην περιγραφή της παρουσιάζεται κάπως έτσι:

Δίνεται η εντολή για την εκφόρτωση του πλοίου οπότε και ο βραχίονας πλησιάζει προς την κατεύθυνση αυτή. Η αρπαγή ανοίγει και πιάνει ένα από τα εμπορευματοκιβώτια (Ε/Κ) και φυσικά το μετακινεί προς τη κατεύθυνση της αποβάθρας. Κάτω από τον γερανό και παραπλεύρως αυτού κινούνται βάσεις τροχήλατες ή σε ράγες, οι οποίες δεν έχουν την ανάγκη οδηγού ή τράκτορα. Αποτελούνται από μικρά κυκλώματα και εγκέφαλο που συνεργάζονται άψογα με το σύστημα του γερανού και είναι προγραμματισμένα να σταματούν σε ορισμένες θέσεις κάτω από τον γερανό έτοιμες να παραλάβουν τα Ε/Κ. Μόλις σταματήσουν και είναι έτοιμες να παραλάβουν το φορτίο τους "κλειδώνουν" στην θέση αυτή και ενεργοποιούν τέσσερις αισθητήρες. Αφού ενεργοποιηθούν αυτοί, ο βραχίονας καθ'επέκτασιν η αρπαγή, παίρνουν την εντολή να κατεβάσουν το Ε/Κ και να το τοποθετήσουν στην βάση αυτή.



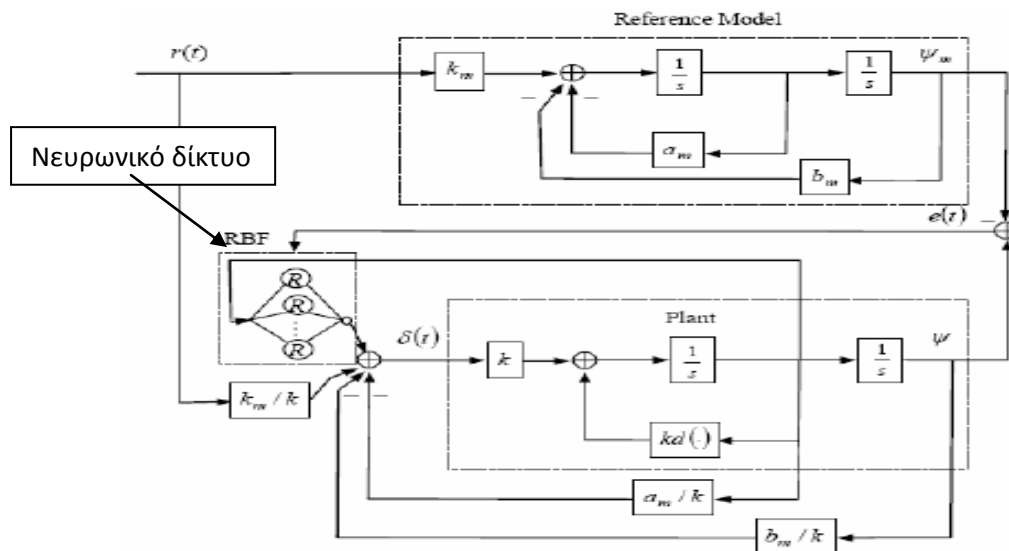
**Εικόνα 8.1:** Διαδικασία φορτοεκφόρτωσης Ε/Κ στο λιμάνι του Rotterdam

Δεν είναι δυνατόν στην φάση αυτή να γίνει κάποιο λάθος στην τοποθέτηση ή να πέσει το Ε/Κ και να προκληθεί ζημιά του φορτιού, μιας και οι αισθητήρες έχουν κι άλλον ένα σημαντικό λόγο ύπαρξης, παραμένουν αναμμένοι και ορίζουν το εμβαδόν τοποθέτησης του φορτιού. Η αρπαγή δεν "βλέπει" τίποτα άλλο εκτός από το ορισμένο εμβαδό και για αυτόν ακριβώς το λόγο μπορεί να το τοποθετήσει μονό εκεί. Διαφορετικά σε περίπτωση βλάβης δεν θα το τοποθετήσει καθόλου έως ότου αποκατασταθεί μέχρι και ο τελευταίος αισθητήρας. Διατηρώντας λοιπόν το φορτίο της, η τροχήλατη ή σε ραγές βάση, απενεργοποιεί τους αισθητήρες της και τότε παίρνει την εντολή να συνεχίσει την πορεία της.

Ο προγραμματισμός ή καλύτερα η εκπαίδευση αυτή έχει πολλά στάδια και λαμβάνονται υπόψη πολλοί παράγοντες και πολλές παράμετροι. Αν για παράδειγμα ένα πλοίο εκφορτώνεται εξ ολοκλήρου και πρέπει τα Ε/Κ να εκφορτωθούν με τέτοιο τρόπο ώστε να φύγουν για τούς μετέπειτα προορισμούς τους γρήγορα, θα πρέπει ίσως να γίνει κάποιου είδους ομαδοποίηση ανά χώρα προορισμού για τον καλύτερο και ευκολότερο χειρισμό τους στην αποβάθρα μετά την εκφόρτωση κ.ο.κ. Όπως πολύ εύκολα μπορεί να καταλάβει κανείς είναι μια αρκετά δύσκολη από "προγραμματιστικής" πλευράς διαδικασία, η οποία απαιτεί συντονισμό και αναλυτική εξέταση των ζητημάτων. Επομένως η ουσία βρίσκεται στις νέες τεχνολογίες που στην περίπτωση του λιμανιού του Rotterdam χρησιμοποιούνται με τον καλύτερο δυνατό τρόπο με σκοπό την προώθηση των συνδυσασμένων εμπορευματικών μεταφορών.

### **8.3: Μοντέλα βασισμένα σε τεχνητά νευρωνικά δίκτυα για τον έλεγχο της διεύθυνσης των πλοίων**

Για να βελτιωθεί η απόδοση του καυσίμου και να μειωθεί η φθορά στα διαφορά εξαρτήματα, χρησιμοποιούνται συστήματα αυτομάτου πιλότου. Συνήθως τα συστήματα αυτομάτου πιλότου βασίζονται σε απλά σχήματα ελέγχου, όπως είναι ο PID έλεγχος. Για την επιτυχέστερη όμως αντιμετώπιση παραμέτρων όπως ο αέρας, τα κύματα, τα θαλάσσια ρεύματα κάτω από διάφορα φορτία, οι ταχύτητες του πλοίου, απαιτούνται πιο εξελιγμένα συστήματα αυτομάτου ελέγχου. Τα τεχνητά νευρωνικά δίκτυα φαίνεται να προσφέρουν πλεονεκτήματα στον τομέα αυτόν εξ αιτίας του πολύ καλού χειρισμού που μπορούν να προσφέρουν σε μη γραμμικά και χαοτικά συστήματα.



Εικόνα 8.2: Σύστημα αυτομάτου ελέγχου για τον έλεγχο της διεύθυνσης του πλοίου χρησιμοποιώντας τεχνητά νευρωνικά δίκτυα

#### **8.4: χρήση τεχνητών νευρωνικών δικτύων στην σταθεροποίηση των πλοίων (απόσβεση ταλαντώσεων).**

Οι Alkan et Al (2004), προτείνουν 2 τεχνητά νευρωνικά δίκτυα για τον καθορισμό της σταθερότητας των αλιευτικών σκαφών. Χρησιμοποιούνται σε κάθε τεχνητό νευρωνικό δίκτυο 2 ενδιάμεσα επίπεδα με 7 και 6 νευρώνες το κάθε επίπεδο αντίστοιχα. Οι εισοδοί του πρώτου επιπέδου από το πρώτο τεχνητό νευρωνικό δίκτυο είναι στοιχεία όπως το βύθισμα και ο λόγος του μήκους προς το εκτόπισμα. Η έξοδος του είναι το κατακόρυφο κέντρο βάρους. Στο δεύτερο τεχνητό νευρωνικό δίκτυο οι εισοδοί είναι στοιχεία όπως το συνολικό μήκος, η πορεία και το βύθισμα στην ίσαλο γραμμή του σκάφους.

#### **8.5: Αυτόματη δημιουργία της μορφής της γάστρας ενός σκάφους.**

Οι Islam et Al (2001) χρησιμοποίησαν τεχνητά νευρωνικά δίκτυα για την σχεδίαση-βελτιστοποίηση της γάστρας ενός σκάφους. Το τεχνητό νευρωνικό δίκτυο είχε 3 επίπεδα και 4 εισόδους, το μήκος, το πλάτος, το σχέδιο και το είδος του πλοίου. Τρεις κρυφές παράμετροι ήταν η ίσαλος γραμμή του πλοίου, η διατομή και η περιοχή στο μέσο του πλοίου. Οι 4 έξοδοι ήταν το εκτόπισμα, το πλάτος, το σχέδιο του πλοίου και η ταχύτητα.

## **8.6: Τεχνητά νευρωνικά δίκτυα για την εύρεση βέλτιστης έλικας πλοίου.**

Έχει παρατηρηθεί πως τα τεχνητά νευρωνικά δίκτυα είναι πολύ αποτελεσματικά εργαλεία για τις εφαρμογές βελτίωσης παραμέτρων. Γι' αυτόν τον λόγο χρησιμοποιούνται συνήθως στην ναυτιλία σε εφαρμογές βελτίωσης κατασκευών των πλοίων. Μια από αυτές είναι και αυτή της εύρεσης βέλτιστης έλικας του πλοίου. Η συγκεκριμένη εφαρμογή ξεπερνάει από πλευράς πολυπλοκότητας τις προηγούμενες, καθώς είναι ένα σύστημα το οποίο περιέχει 555 ρυθμίσεις παραμέτρων εισόδων και επιθυμητών εξόδων στο δίκτυο.

## Επίλογος – Συμπεράσματα

- Από την μέχρι τώρα εμπειρία με τα δίκτυα, είναι προφανές πως η χορήγηση σε αυτά δεδομένων διασταύρωσης επιφέρει μονό θετικά αποτελέσματα αφού εκτός της καλύτερης προσέγγισης, συνεισφέρουν στην ταχύτερη και καλύτερη εκπαίδευση του δικτύου. Μειώνουν έτσι τις υπολογιστικές ώρες που απαιτούνται και αναβαθμίζουν ποσοτικά την έρευνα.
- Η χρήση περισσότερων δεδομένων θα οδηγήσει πιθανότατα σε καλύτερα αποτελέσματα. Σε αυτό το συμπέρασμα καταλήγουμε αν συγκρίνουμε τα αποτελέσματα των προηγούμενων παραδειγμάτων με τα εκάστοτε δεδομένα εισαγωγής.
- Επίσης επιβεβαιώνεται πως τα Τεχνητά Νευρωνικά Δίκτυα μπορούν να χρησιμοποιηθούν ως εργαλεία πρόβλεψης και να παράγουν αξιόπιστα αποτελέσματα για το μέλλον.
- Από την παρούσα εργασία αλλά και από τις προγενέστερες αυτής μπορούμε να συμπεράνουμε πως στη ναυτιλία υπάρχει πρόσφορο έδαφος για την εφαρμογή των δικτύων αφού όλα τα στοιχεία οδηγούν προς αυτή την κατεύθυνση.

## Βιβλιογραφία

1. Σχιζά Α (2007) - *Σχέση Ναυτιλιακών Παραγώγων και Εξέλιξης Επιπέδου Ναύλων*
2. Ρίζος Γ. (1996) - *Τεχνητά Νευρωνικά Δίκτυα, Εκδόσεις Νέων Τεχνολογιών*
3. J. A. Anderson (1995) - *An Introduction to Neural Networks*, MIT Press, Cambridge (1995)
4. J. E. Dayhoff (1990) - *Neural Network Architectures, an Introduction*, Van Nostrand Reinhold (New York), 1990
5. P. Picton (1994) - *Introduction to Neural Networks*, Macmillan, London
6. Αργυράκης Π. (2001) - *Νευρωνικά δίκτυα και εφαρμογές*
7. Halbert White (1996) - *Economic Prediction using Neural Networks: The Case of IBM Daily Stock Returns*
8. ΧΑΡΙΛΑΟΣ Ν ΨΑΡΑΥΤΗΣ (1999) - *Συστήματα Θαλασσιών Μεταφορών*
9. Neural networks: A Requirement for Intelligent Systems, <http://www.learnartificialneuralnetworks.com/>
10. Stopford M. (2009) - *Maritime Economics 3rd edition*, Routledge, London.
11. Lippmann Richard P. (1987), *An Introduction to Computing with Neural Nets*
12. Γλυνός Νικόλαος, Γεωργίου Β. Χάρης, μελισσόβας Β. Σπύρος, Παπαδόπουλος Σ. Δημήτρης (1995) – *Μέθοδοι Εκπαίδευσης και Μοντέλα Τεχνικών Νευρωνικών Δικτύων*, Πανεπιστήμιο Ιωαννίνων, Τμήμα Πληροφορικής.
13. Διαμαντάρας Κωνσταντίνος (2007), *Τεχνητά Νευρωνικά Δίκτυα*, Εκδόσεις Κλειδάριθμος, Αθήνα.
14. Omri Weisman and Ziv Pollak (1995), *The Perceptron*, Ben Gurion University of the Negev.
15. Carling, A. (1992). *Introducing Neural Networks*. Wilmslow, UK: Sigma Press
16. *International Journal of Information Technology*, Vol. 11 No. 6 (2005)

## Περιεχόμενα

Περίληψη .....	3
Abstract.....	4
Πρόλογος .....	5
Κεφάλαιο 1: ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ .....	6
/Κεφάλαιο 2: ΤΙ ΕΙΝΑΙ ΕΝΑ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ .....	9
Κεφάλαιο 3: ΔΟΜΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ .....	12
Κεφάλαιο 4: ΜΑΘΗΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ.....	14
Κεφάλαιο 5: ΑΞΙΟΠΙΣΤΙΑ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ .....	16
Κεφάλαιο 6: ΠΡΟΤΥΠΑ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ.....	20
Κεφάλαιο 7: ΕΦΑΡΜΟΓΕΣ .....	25
Κεφάλαιο 8: ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΝΑΥΤΙΛΙΑ .....	27
Επίλογος – Συμπεράσματα .....	31
Βιβλιογραφία .....	32