



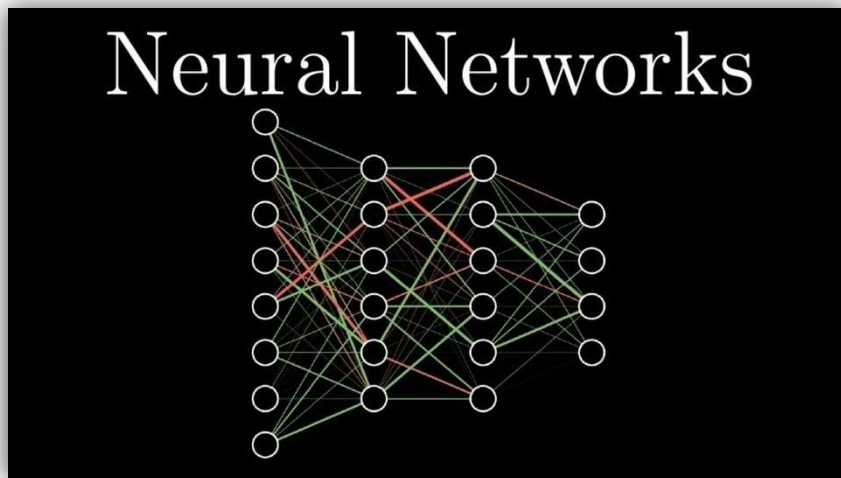
ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΑΤΡΩΝ  
UNIVERSITY OF PATRAS

ΣΧΟΛΗ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ

ΤΜΗΜΑ ΔΙΟΙΚΗΣΗΣ ΕΠΙΧΕΙΡΗΣΕΩΝ ΠΑΤΡΑΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

# «NEURAL NETWORKS AND LEARNING MACHINES»



Όνοματεπώνυμο: Ηλίας Καταβάτης

Όνοματεπώνυμο: Ευάγγελος Αντώνιος Γάκης

Επιβλέπων Καθηγητής: Δημήτρης Παπαδόπουλος

Πάτρα, 2019

# ΕΥΧΑΡΙΣΤΙΕΣ

---

Θα θέλαμε να ευχαριστήσουμε τον καθηγητή μας για την εργασία που μας ανέθεσε με θέμα τα Τεχνητά Νευρωνικά Δίκτυα και τη Μηχανική Μάθηση. Καθώς και την καθοδήγηση του και για τις χρήσιμες συμβουλές που μας έδωσε με σκοπό να φέρουμε εις πέρας την πτυχιακή μας εργασία.

# ΠΕΡΙΛΗΨΗ

---

Στην πτυχιακή αυτή εργασία πραγματοποιείται μια ανάλυση του κλάδου των Νευρωνικών Δικτύων (Neural Networks) και της Μηχανικής Μάθησης (Learning Machine). Αποτελεί έναν επιστημονικό κλάδο, ο οποίος αναπτύχθηκε ραγδαία τις τελευταίες δεκαετίες.

Πρόκειται για ένα σύνολο διασυνδεδεμένων μονάδων, οι οποίες δημιουργούν ένα μοντέλο, με σκοπό να εκτελεί διάφορες λειτουργίες. Τα Τεχνητά Νευρωνικά Δίκτυα χρησιμοποιούνται σε πολλούς και διαφορετικούς επιστημονικούς κλάδους. Ο κύριος στόχος για τη δημιουργία και την ανάπτυξη του ήταν η προσομοίωση των πολύπλοκων λειτουργιών που εκτελούν τα Βιολογικά Νευρωνικά δίκτυα όπως ο ανθρώπινος εγκέφαλος.

Η συγκεκριμένη εργασία αποτελείται συνολικά από έξι κεφάλαια.

Πιο αναλυτικά, στο πρώτο κεφάλαιο πραγματοποιείται μια εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα. Αναφερόμαστε στη λειτουργία των Τεχνητών Νευρωνικών Δικτύων καθώς και στην αντιστοιχία τους με τα Βιολογικά Νευρωνικά Δίκτυα. Αναφέρονται και αναλύονται περαιτέρω οι αρχιτεκτονικές των Δικτύων αυτών καθώς και τα πλεονεκτήματα που διαθέτουν. Στο τέλος του συγκεκριμένου κεφαλαίου αναφέρονται ορισμένοι κλάδοι στους οποίους βρίσκουν εφαρμογή τα Τεχνητά Νευρωνικά Δίκτυα.

Στο δεύτερο κεφάλαιο αυτής της εργασίας γίνεται αναφορά στις βασικές έννοιες της Μηχανικής Μάθησης καθώς και στους τρόπους με τους οποίους εκπαιδεύονται τα Τεχνητά Νευρωνικά Δίκτυα. Τέλος, περιγράφονται τα είδη και οι λειτουργίες της Μηχανικής Μάθησης.

Το τρίτο κεφάλαιο ασχολήθηκε αποκλειστικά με την ανάλυση του απλού δικτύου Perceptron, το οποίο αποτελείται από έναν και μόνο νευρώνα. Αναλύονται οι τρόποι εκπαίδευσης του συγκεκριμένου αισθητήρα. Επιπλέον, αναφέρεται το μειονέκτημα που υπάρχει στο δίκτυο Perceptron καθώς και μια περιγραφή του προβλήματος της συνάρτησης XOR.

Μετά την ανάλυση του απλού δικτύου Perceptron, στο τέταρτο κεφάλαιο αναφερόμαστε στο δίκτυο Perceptron που αποτελείται από πολλά επίπεδα. Αναφέρονται οι τρόποι εκπαίδευσης του συγκεκριμένου δικτύου.

Πραγματοποιείται μια αναφορά στον αλγόριθμο Error Back Propagation καθώς και το μειονέκτημα που έχει ο συγκεκριμένος αλγόριθμος.

Το πέμπτο κεφάλαιο αναφέρεται στον αλγόριθμο του μέσου τετραγωνικού σφάλματος Least Mean Square – LMS καθώς και στη διαδικασία φιλτραρίσματος του συγκεκριμένου αλγορίθμου.

Στο έκτο και τελευταίο κεφάλαιο της συγκεκριμένης εργασίας πραγματοποιείται μια περιγραφή της μελλοντικής εξέλιξης του κλάδου των Νευρωνικών Δικτύων καθώς και τα τελικά συμπεράσματα που προέκυψαν ύστερα από τη μελέτη του συγκεκριμένου επιστημονικού κλάδου.

# ABSTRACT

---

In this thesis an analysis of sector of the Neural Networks and the Learning Machine is carried out. This sector has grown up rapidly in the last decades.

It's about a total of interconnected units, that create a model that performs various functions. The Artificial Neural Networks are used in many different scientific disciplines. The main goal of its creation and development was to simulate the complex functions performed by Biological Neural Networks such as the human brain.

This thesis consists of six chapters.

More specifically, in the first chapter, an introduction of the sector of Artificial Neural Networks is carried out. We are referring to the function of Artificial Neural Networks, as well as to the connection among them and the Biological Neural Networks. We are referring to the architectures of the Neural Networks and the advantages that they have. At the end of this chapter, we are going to analyze the sectors that the Neural Networks use.

In the second chapter of this thesis, an introduction of Machine Learning is carried out. We are referring to the learning processes of Machine Learning, which the Neural Networks use. Finally, we describe the types and the functions of the Machine Learning.

The third chapter dealt exclusively with the analysis of the Rosenblatt's Perceptron network, which consists of a single neuron. It analyzes the training of the particular network. In addition, describe the disadvantage that exist in the Perceptron network and problem of the XOR function that are referred.

In the fourth chapter we are referring to the Multilayer Perceptron consisting of several layers. And also describe all the different ways of training this neural network. A description of the Error Back Propagation algorithm is carried out in this chapter, too. In addition, the disadvantage of the particular algorithm is referred.

The fifth chapter refers to the Least Mean Square - LMS algorithm. Also refers to the filtering structure of this particular algorithm.

In the sixth and the final chapter, a description of the future evolution and the conclusions of the Neural Networks sector is carried out.

# ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ .....	1
ΠΕΡΙΛΗΨΗ .....	2
ABSTRACT .....	4
ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ .....	8
ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	9
1.1    Τι είναι τα Νευρωνικά Δίκτυα.....	9
1.2    Ιστορική αναδρομή .....	10
1.3    Η λειτουργία των νευρωνικών δικτύων.....	12
1.4    Νευρωνικά Δίκτυα ως κατευθυνόμενοι γράφοι.....	14
1.5    Μοντέλο τεχνητού νευρώνα .....	16
1.6    Ικανότητα γενίκευσης .....	21
1.7    Βιολογικά Νευρωνικά Δίκτυα.....	22
1.7.1    Εγκέφαλος: ένα σύνθετο βιολογικό νευρωνικό δίκτυο .....	23
1.7.2    Βιολογικός Νευρώνας .....	24
1.7.3    Υπολογιστής και Ανθρώπινος Εγκέφαλος .....	26
1.8    Αρχιτεκτονικές νευρωνικών δικτύων.....	26
1.8.1    Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης (feedforward) .....	26
1.8.2    Νευρωνικά Δίκτυα με ανατροφοδότηση (feedback, recurrent).....	27
1.9    Πλεονεκτήματα .....	28
1.10   Σύγχρονες Εφαρμογές.....	29
1.11   Βασικές τεχνικές και εφαρμογές .....	31
1.11.1   Interpolation .....	31
1.11.2   Classification.....	35
1.11.3   Inverse problems .....	38
1.11.4   Control applications.....	40
ΚΕΦΑΛΑΙΟ 2 – ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ .....	43
2.1    Εισαγωγή – Βασικές έννοιες .....	43
2.2    Εκπαίδευση Νευρωνικών Δικτύων.....	44

2.3	Είδη Μηχανικής Μάθησης .....	45
2.3.1	Επιβλεπόμενη Μάθηση (Supervised Learning) .....	45
2.3.2	Ενισχυτική Μάθηση (Reinforcement Learning) .....	46
2.3.3	Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning) .....	48
2.4	Λειτουργίες Μάθησης .....	49
2.4.1	Συσχέτιση προτύπων .....	49
2.4.2	Αναγνώριση προτύπων (pattern recognition) .....	50
2.4.3	Έλεγχος .....	51
2.5	Αλγόριθμοι Μάθησης .....	51
ΚΕΦΑΛΑΙΟ 3 -ΝΕΥΡΩΝΑΣ PERCEPTRON .....		53
3.1	Εισαγωγή .....	53
3.2	Δίκτυο Perceptron .....	53
3.3	Ταξινόμηση προτύπων .....	57
3.4	Η εκπαίδευση του αισθητήρα .....	60
3.5	Αλγόριθμος σύγκλισης .....	61
3.6	Μειονέκτημα του δικτύου Perceptron .....	63
3.7	Πρόβλημα Συνάρτησης XOR .....	63
ΚΕΦΑΛΑΙΟ 4 – PERCEPTRON ΠΟΛΛΑΠΛΩΝ ΕΠΙΠΕΔΩΝ .....		66
4.1	Πολυεπίπεδο δίκτυο Perceptron – MLP .....	66
4.2	Αλγόριθμος Error Back Propagation .....	71
4.3	Μειονεκτήματα αλγορίθμου οπισθοδιάδοσης .....	73
4.4	Τρόποι εκπαίδευσης δικτύου MPL .....	74
ΚΕΦΑΛΑΙΟ 5 – ΑΛΓΟΡΙΘΜΟΣ ΜΕΣΟΥ ΤΕΤΡΑΓΩΝΙΚΟΥ ΣΦΑΛΜΑΤΟΣ .....		77
5.1	Εισαγωγή .....	77
5.2	Διαδικασία Φιλτραρίσματος του LMS αλγορίθμου .....	78
5.3	Βελτιστοποίηση χωρίς περιορισμούς .....	83
5.4	Αλγόριθμος Least Mean Square - LMS .....	84
ΚΕΦΑΛΑΙΟ 6 – ΣΥΜΠΕΡΑΣΜΑΤΑ .....		87
6.1	Συμπεράσματα .....	87
6.2	Μελλοντικές εξελίξεις των Νευρωνικών Δικτύων .....	89
ΒΙΒΛΙΟΓΡΑΦΙΑ .....		91



# ΠΙΝΑΚΑΣ ΕΙΚΟΝΩΝ

---

Εικόνα 1 - Συναπτική σύνδεση .....	15
Εικόνα 2 - Συνδέσεις ενεργοποίησης.....	15
Εικόνα 3 - Απεικόνιση σήματος κόμβου.....	16
Εικόνα 4 - Αρχιτεκτονικός Γράφος .....	16
Εικόνα 5 - Μοντέλο Τεχνητού Νευρώνα.....	17
Εικόνα 6 - Συνάρτηση Ενεργοποίησης Κατωφλιού.....	19
Εικόνα 7 - Σιγμοειδής Συνάρτηση Ενεργοποίησης.....	20
Εικόνα 8 - Βιολογικός Νευρώνας .....	25
Εικόνα 9 - Τεχνητό Νευρωνικό Δίκτυο πρόσθιας τροφοδότησης.....	27
Εικόνα 10 - Τεχνητό Νευρωνικό Δίκτυο Ανατροφοδότησης.....	28
Εικόνα 11 - Δομή νευρωνικού δικτύου .....	33
Εικόνα 12 - Σύγκριση αποτελεσμάτων για νευρωνικό δίκτυο με την παλινδρόμηση.....	34
Εικόνα 13 - Τροποποιημένη έκδοση του δικτύου.....	35
Εικόνα 14 - Σχηματική διατομή ενός αγωγού πετρελαίου .....	36
Εικόνα 15 - Τέσσερις διαμορφώσεις μοντέλων τριών φάσεων ροής.....	37
Εικόνα 16 - Το πρόβλημα της τομογραφίας .....	39
Εικόνα 17 - Επιτυχής έλεγχος ανατροφοδότησης του tokamak.....	41
Εικόνα 18 - Διάγραμμα επιβλεπόμενης μάθησης (Supervised Learning).....	45
Εικόνα 19 - Ενισχυτική Μάθηση (Reinforcement Learning).....	47
Εικόνα 20 - Διάγραμμα μη Επιβλεπόμενη Μάθηση (Unsupervised Learning) .....	49
Εικόνα 21 - Αισθητήρας Perceptron .....	54
Εικόνα 22 - Αναπαράσταση του υπερεπιπέδου για την ταξινόμηση δύο κλάσεων....	59
Εικόνα 23 - Πρόβλημα XOR.....	64
Εικόνα 24 - Πολυεπίπεδο δίκτυο Perceptron .....	66
Εικόνα 25 - Απεικόνιση των σημάτων ενός MLP δικτύου .....	69
Εικόνα 26 - Δυναμικό Σύστημα.....	79
Εικόνα 27 - Γράφημα ροής σήματος προσαρμοστικού μοντέλου .....	80
Εικόνα 28 - Ροή Σήματος του αλγορίθμου LMS.....	85

# ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

---

## 1.1 Τι είναι τα Νευρωνικά Δίκτυα

Με τον όρο «Νευρωνικά Δίκτυα» αναφερόμαστε σε ένα κύκλωμα διασυνδεδεμένων υπολογιστικών κόμβων, τους οποίους ονομάζουμε νευρώνες. Διαχωρίζονται σε βιολογικούς και τεχνητούς νευρώνες. Πιο συγκεκριμένα, στους βιολογικούς αναφερόμαστε σε ένα μέρος τους νευρικού ιστού, ενώ στους τεχνητούς αναφερόμαστε σε αλγοριθμικό δημιούργημα. (Haykin, 2009)

Τα Νευρωνικά Δίκτυα συνηθίζεται να αναφέρονται και ως Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks, ANN). (Haykin, 2009) Αποτελεί ένα είδος τεχνολογίας εκμάθησης που εμπίπτει στον τομέα της Τεχνητής Νοημοσύνης. Πρόκειται για έναν νέο τομέα των φυσικών επιστημών, ο οποίος συνεχώς αναπτύσσεται και εξελίσσεται. Τα Νευρωνικά δίκτυα τα τελευταία χρόνια έχουν παρουσιάσει μεγάλη ανάπτυξη, κεντρίζοντας το ενδιαφέρον πολλών επιστημών. Ο λόγος ήταν γιατί θεώρησαν πολύ σημαντική και χρήσιμη τη δημιουργία και κατασκευή είτε μιας μηχανής είτε ενός προγράμματος τα οποία θα εκτελούσαν διεργασίες που μόνο ο ανθρώπινος εγκέφαλος μπορούσε να εκτελέσει.

Το βασικότερο χαρακτηριστικό των τεχνητών νευρωνικών δικτύων από τα πρώτα κιάλας πρότυπα που αναπτύχθηκαν ήταν ότι οι διεργασίες τους βασίστηκαν στο νευρικό σύστημα του ανθρώπου. Με το πέρασμα των χρόνων και τη ραγδαία ανάπτυξη τους τα νευρωνικά δίκτυα έχουν εξελιχθεί πέρα από τους βιολογικούς οργανισμούς. Σήμερα χρησιμοποιούνται με σκοπό να επιλύουν πολλών ειδών προβλήματα με χρήση ηλεκτρονικού υπολογιστή. Μια σημαντική διαφορά που υπάρχει ανάμεσα στα δίκτυα αυτά και τους κλασικούς υπολογιστές είναι ως προς τη λειτουργία τους. Ο τρόπος λειτουργίας των δικτύων είναι να συνδυάζουν αρμονικά ένα μαθηματικό τρόπο σκέψης με τον ανθρώπινο τρόπο σκέψης. Πιο αναλυτικά, μπορούμε να αναφέρουμε ότι ένα δίκτυο έχει τη δυνατότητα να εκπαιδεύεται και να μαθαίνει, να απομνημονεύει μια τιμή, γεγονός που εμπεριέχεται στον ανθρώπινο τρόπο σκέψης. Έχουν τη δυνατότητα επίσης να χρησιμοποιούν κάθε είδους περίπλοκη μαθηματική συνάρτηση. (Αργυράκης, 2001)

Αξιοσημείωτο είναι το γεγονός ότι με τον τομέα των νευρωνικών δικτύων ασχολούνται επιστήμονες διαφόρων φυσικών επιστημών όπως, λόγου χάρη τη φυσική, τα μαθηματικά, την ιατρική, την επιστήμη των υπολογιστών, την επιστήμη των μηχανικών κλπ. Αυτό συμβαίνει διότι για τα νευρωνικά δίκτυα

απαιτούνται γνώσεις από διάφορους επιστημονικούς κλάδους, προκειμένου να υλοποιηθούν οι τεχνικές και οι μέθοδοι που απαιτούνται. Σε καμία άλλη επιστήμη στις μέρες μας δεν παρατηρείται τόσο άμεσος συνδυασμός από επιστημονικές γνώσεις πολλών διαφορετικών πεδίων.

Τα τεχνητά νευρωνικά δίκτυα εκπαιδεύονται με σκοπό να παρουσιάσουν την καλύτερη και πιο εφικτή γενικευτική ιδιότητα, όταν καλούνται να αναγνωρίσουν νέα δεδομένα. Αυτό αποτελεί κύριο στόχο στη δημιουργία των νευρωνικών δικτύων, να κατασκευαστεί δηλαδή η εσωτερική τους δομή κατάλληλα ώστε να έχουν τη δυνατότητα να αναγνωρίζουν άγνωστα μέχρι στιγμής πρότυπα, τα οποία ταιριάζουν με άλλα πρότυπα με τα οποία εκπαιδεύτηκαν. (Αργυράκης, 2001)

Ένας ορισμός που μπορεί να δοθεί για τα νευρωνικά δίκτυα είναι ο ακόλουθος:

Νευρωνικό δίκτυο ονομάζετε ένας επεξεργαστής, ο οποίος αποτελείται από ένα πλήθος απλών μονάδων για την αποθήκευση των γνώσεων και την εκτέλεση διάφορων λειτουργιών. (Haykin, 2009)

## 1.2 Ιστορική αναδρομή

Ο κλάδος των νευρωνικών δικτύων είναι ένας σχετικά καινούργιος κλάδος της τεχνολογίας. Η ανάπτυξη του πέρασε από πολλά στάδια και η εξέλιξή του ήταν ραγδαία. Ξεκίνησε μόλις τις τελευταίες δεκαετίες αλλά η μεγάλη ώθηση πραγματοποιήθηκε μετά το 1980. (Αργυράκης, 2001)

Το 1943 παρουσιάστηκε το πρώτο μοντέλο νευρωνικού δικτύου από τον νευροφυσιολόγο McCulloch και τον 18χρονο πρωτοετή φοιτητή Μαθηματικών, Pitts . Το μοντέλο αυτό υποστήριξε ότι οι νευρώνες αποτελούν τη βασική μονάδα ενός δικτύου. Υποστήριξαν επίσης για πρώτη φορά ότι ένα νευρωνικό δίκτυο είναι ένα σύνολο από νευρώνες συνδεδεμένοι μεταξύ τους. Οι δύο συγγραφείς θεώρησαν ότι οι νευρώνες και οι διασυνδέσεις τους αποτελούν ένα πρότυπο, παρόμοιο με ένα ηλεκτρικό κύκλωμα. Το 1947 παρουσίασαν ένα εξελιγμένο πρότυπο προκειμένου να πραγματοποιείται αναγνώριση σχημάτων. (D. Anderson, 1992)

Ο J. von Neumann βασίστηκε στις εργασίες που υποστήριξαν οι McCulloch και Pitts για τις υπολογιστικές μηχανές την δεκαετία του πενήντα. Εκείνη τη χρονική περίοδο πραγματοποιήθηκαν οι πρώτες προσπάθειες να βασιστούν στα βιολογικά δίκτυα και να αντλήσουν πληροφορίες από αυτά με σκοπό την ανάπτυξη των πρώτων τεχνητών δικτύων. (Αργυράκης, 2001)

Ο συγγραφέας Donald Hebb το 1949 έγραψε το βιβλίο του με τίτλο “The organization of behavior”. (D. Anderson, 1992) Στο βιβλίο αυτό, ο συγγραφέας εισάγει για πρώτη φορά το μοντέλο μάθησης, το οποίο προέκυψε ύστερα από πληθώρα πειραμάτων νευροφυσιολογίας. Το συγκεκριμένο μοντέλο μάθησης βασίζεται στην ιδέα των διασυνδέσεων ανάμεσα στις μονάδες (νευρώνες) που υπάρχουν σε οποιοδήποτε σύστημα. Πιο συγκεκριμένα αναφέρει ότι κάθε φορά που οι νευρωνικές συνδέσεις χρησιμοποιούνται από το ίδιο το δίκτυο, το αποτέλεσμα είναι ότι αυτές οι συνδέσεις ενισχύονται περισσότερο. Το γεγονός αυτό έχει ως συνέπεια το δίκτυο κάθε φορά να πλησιάζει περισσότερο στα επιθυμητά αποτελέσματα, να εκπαιδευτεί δηλαδή στο πρότυπο το οποίο παρουσιάζεται.

Το 1957 ο F. Rosenblatt παρουσίασε για πρώτη φορά το μοντέλο του Αισθητήρα - Perceptron. Ο F. Rosenblatt είχε κατασκευάσει αρχικά ένα δίκτυο με Υλικό (Hardware). Το δίκτυο αυτό εκτελούσε αρκετές διεργασίες. Πρόκειται για ένα σχετικά απλό μοντέλο δύο επιπέδων, εισόδου και εξόδου, με μονοδρομική κατεύθυνση του σήματος από την είσοδο προς την έξοδο. Το μοντέλο του Αισθητήρα τότε είχε παρουσιάσει πολλές επιτυχίες και δεν ήταν λίγοι όσοι πίστεψαν ότι τα νευρωνικά δίκτυα είναι μια ανώτερη τεχνική επίλυσης άλυτων προβλημάτων. (Αργυράκης, 2001) Παρά την μεγάλη επιτυχία του μοντέλου Perceptron, παρουσιάστηκαν αρκετοί περιορισμοί.

Το 1969, οι Minsky και Papert παρουσίασαν στο βιβλίο τους με τίτλο «Perceptrons», στο οποίο αναλύθηκε μια συνολική, σφαιρική εικόνα του προτύπου. Στο βιβλίο δίνεται έμφαση στην σπουδαιότητα του προτύπου καθώς και στο σύνολο των ενεργειών, τις οποίες μπορεί να πραγματοποιήσει. (D. Anderson, 1992). Αποδεικνύεται εμπειριστικώς ότι στα συγκεκριμένα πρότυπα παρουσιάζονται αρκετοί περιορισμοί. Για το λόγο αυτό, ο κλάδος των νευρωνικών δικτύων εκείνη την χρονική περίοδο χάνει τη δημοτικότητα του. Η προσοχή της κοινωνίας στρέφεται σε ένα παρεμφερή κλάδο, ο οποίος ήταν ο κλάδος της Τεχνητής Νοημοσύνης.

Το 1959 οι Bernard Widrow και Marcian Hoff ανέπτυξαν δύο καινούργια πρότυπα του αισθητήρα. Τα μοντέλα αυτά ονομάστηκαν Adaline και Madaline και αποτέλεσαν τα πρώτα μοντέλα, τα οποία χρησιμοποιήθηκαν για διάφορα πρακτικά προβλήματα με μεγάλη επιτυχία (D. Anderson, 1992). Ένα τέτοιο πρόβλημα που επιλύθηκε επιτυχώς ήταν η χρήση τους ως φίλτρα προκειμένου να εξαλειφθεί η ηχώ από τις τηλεφωνικές γραμμές.

Το 1982 ο βιολόγος J. Hopfield παρουσίασε ένα έργο, το οποίο έδωσε μεγάλο έναυσμα στην εξέλιξη των δικτύων. Η έκταση της εργασίας του ήταν μόλις λίγες σελίδες και αποδείκνυε μέσα από αυτές με μαθηματικές αποδείξεις ότι τα νευρωνικά δίκτυα έχουν την ικανότητα να χρησιμοποιηθούν ως αποθηκευτικοί χώροι καθώς και ότι μπορούν να επανακτούν τις πληροφορίες του συστήματος (Αργυράκης, 2001) Η εργασία του J. Hopfield αποτέλεσε έμπνευση για μελλοντικές ιδέες και εξελίξεις του συγκεκριμένου κλάδου.

Το 1986 οι McClelland και Rumelhart παρουσίασαν ένα ακόμα ιδιαίτερα σημαντικό έργο με τίτλο “Parallel Distributed Processing” που σχετίζεται με την εκπαίδευση του νευρωνικού δικτύου. Βασίζετε στην ιδέα ότι τα δίκτυα μπορούν να χρησιμοποιηθούν ως ένας παράλληλος επεξεργαστής. Το έργο των McClelland και Rumelhart έχει συμβάλει σε μεγάλο βαθμό στην εξέλιξη των νευρωνικών δικτύων καθώς έχουν προχωρήσει πέραν του αισθητήρα (Perceptron). Πιο συγκεκριμένα, έκαναν εφικτή την ύπαρξη επιπλέον επιπέδων νευρώνων, πέραν από την βασική εσωτερική δομή των δικτύων που αποτελούσε την είσοδο και την έξοδο. Η τεχνική της οπισθοδιάδοσης (back-propagation), η οποία χρησιμοποιείτε και στις μέρες μας είναι μια μέθοδος εκπαίδευσης δικτύων που πρότειναν οι McClelland και Rumelhart στο συγκεκριμένο έργο τους. (P.J. Braspenning, A.J.M.M. Weijters, F. Thuijsman, 1995 )

Τα πρώτα συνέδρια που ασχολήθηκα αποκλειστικά με τα νευρωνικά δίκτυα πραγματοποιήθηκαν το 1985 από τους American Physical Society και IEEE. Την ίδια χρονική περίοδο ξεκίνησαν να δημιουργούνται διάφορες εταιρίες που δραστηριοποιούνται αποκλειστικά με τα νευρωνικά δίκτυα. Ένα αντιπροσωπευτικό παράδειγμα τέτοιας εταιρίας είναι η International Neural Network Society, η οποία αναπτύχθηκε στην Αμερική, στην Ευρώπη και στην Ιαπωνία.

Τα χρόνια που πέρασαν και συγκεκριμένα τη δεκαετία του 90, παρατηρήθηκε ότι κλάδος των νευρωνικών δικτύων είχε αναπτυχθεί σε πολύ μεγάλο βαθμό και πλήθος από επιστήμονες ασχολούνται αποκλειστικά με αυτόν τον τομέα. Αποτελούσε πλέον ένα ανεξάρτητο επιστημονικό πεδίο. Αναπτύχθηκαν τουλάχιστον 10 περιοδικά επιστήμης, που αφιερώθηκαν στα Τεχνητά Νευρωνικά Δίκτυα. (Αργυράκης, 2001)Τη συγκεκριμένη δεκαετία επίσης πραγματοποιήθηκαν πολλά ετήσια συνέδρια με πολλούς παραβρισκόμενους, που ασχολήθηκαν αποκλειστικά με την ανάλυση αυτών των δικτύων. Επιπλέον, δημιουργήθηκαν πολλές εμπορικές εφαρμογές γύρω από τον κλάδο αυτό, κυρίως από τις Ηνωμένες Πολιτείες της Αμερικής.

Μέχρι σήμερα ο κλάδος των νευρωνικών δικτύων παρουσίασε σημαντική εξέλιξη. Παρόλα αυτά, στις μέρες μας παρατηρείται ένας μειωμένος ρυθμός ανάπτυξης και εξέλιξης των τεχνητών νευρωνικών δικτύων.

### **1.3 Η λειτουργία των νευρωνικών δικτύων**

Ένα δίκτυο αποτελείται από έναν μεγάλο αριθμό επεξεργαστών - μονάδων διατεταγμένοι σε διάφορα επίπεδα, οι οποίοι εκτελούν τις λειτουργίες τους ταυτόχρονα. Οι επεξεργαστές που βρίσκονται στην πρώτη βαθμίδα δέχονται τις πρώτες πληροφορίες εισόδου. Κάθε επίπεδο δέχεται την έξοδο από το ακριβώς προηγούμενο επίπεδο του. Το τελευταίο επίπεδο λαμβάνει εισόδους από το επίπεδο που προηγείται και παράγει την έξοδο του συστήματος.

Προκειμένου να κατασκευαστεί ένα νευρωνικό δίκτυο, πρώτο μέλημα είναι ο ορισμός της τοπολογίας του. Με την τοπολογία ενός δικτύου αναφερόμαστε στο πλήθος των νευρώνων που υπάρχουν σε αυτό καθώς και την δομή των συνάψεων, με τα τις οποίες μεταδίδονται τα δεδομένα ανάμεσα στους νευρώνες. (Πλέρου, 2012)

Όπως ήδη έχει αναφερθεί, οι νευρώνες ενός δικτύου ομαδοποιούνται σε διάφορα επίπεδα (layers). Ένας νευρώνας αναφέρεται ως εσωτερικός ή κρυμμένος (hidden) όταν δεν διαθέτει καμία σύνδεση με το εξωτερικό περιβάλλον. Επίσης, είναι ιδιαίτερα σημαντικό να αναφερθεί ότι η τοπολογία ενός τεχνητού νευρωνικού δικτύου διακρίνεται σε δύο μορφές, δυναμική (μεταβαλλόμενη) ή στατική. (Haykin, 2009)

Κάθε μονάδα εκτελεί μια συγκεκριμένη συνάρτηση, με σκοπό το δίκτυο να καταφέρει με τη σειρά του να επιτύχει μια συγκεκριμένη διεργασία. Με τη επιλογή της σωστής μεθόδου εκπαίδευσης, καθορίζονται οι τιμές των βαρών του δικτύου, προκειμένου να επιτευχθούν οι λειτουργίες του. Η εκπαίδευση ενός τεχνητού νευρωνικού δικτύου καθώς και η εμπειρία και η γνώση που αποκτά βρίσκονται αποθηκευμένα στις τιμές των συνδέσεων των νευρώνων. (Haykin, 2009)

Οι νευρώνες αποτελούν τα βασικά, δομικά στοιχεία ενός δικτύου. Σε κάθε κόμβο εισέρχεται ένα σύνολο εισόδων από διάφορες πηγές. Οι πηγές αυτές μπορεί να είναι είτε άλλοι νευρώνες, είτε το περιβάλλον. Με βάση τις συγκεκριμένες τιμές που δέχεται ως εισόδους, επιτελεί έναν υπολογισμό και παράγεται μια έξοδος. Η παραγόμενη έξοδος κάθε νευρώνα μπορεί είτε να αποτελεί είσοδο ενός άλλου νευρώνα μέσα στο δίκτυο, είτε να κατευθύνεται στο περιβάλλον.

Οι τεχνητοί νευρώνες ενός δικτύου αλληλεπιδρούν διαρκώς μεταξύ τους και είναι συνδεδεμένοι με τις λεγόμενες συνάψεις (synapses). Για κάθε ζευγάρι νευρώνων διαφέρει το μέγεθος της αλληλεπίδρασης αλλά προσδιορίζεται από τα συναπτικά βάρη (synaptic weights). Πιο συγκεκριμένα, τα συναπτικά βάρη συνεχώς μεταβάλλονται όσο το δίκτυο στο οποίο ανήκουν αλληλεπιδρά με το εξωτερικό περιβάλλον. (Haykin, 2009) Για το λόγο αυτό, διαρκώς ενδυναμώνεται ή αποδυναμώνεται η ισχύς των δεσμών του δικτύου. Επομένως, η λειτουργία που επιτελούν τα συναπτικά βάρη είναι ιδιαίτερα σημαντική, διότι η γνώση που αποκτά το δίκτυο από το εξωτερικό του περιβάλλον κωδικοποιείται σε αυτά. Έτσι, κάθε νευρωνικό δίκτυο αποκτά την ικανότητα για προσαρμογή στο περιβάλλον και συνεχή εξέλιξη.

Ένα δίκτυο μπορεί να εκπαιδευτεί με δύο τρόπους. Σύμφωνα με τον πρώτο τρόπο η εκπαίδευση πραγματοποιείται με εποπτεία. Πιο συγκεκριμένα, κάθε δίκτυο που εκπαιδεύεται με εποπτεία δέχεται ένα σύνολο καταστάσεων στις οποίες μπορεί να φτάσει καθώς και τα επιθυμητά αποτελέσματα. Με σκοπό να εκπαιδύσουμε το δίκτυο να γνωρίσει τα παραδείγματα αυτά χρειάζεται να χρησιμοποιηθεί ένας αλγόριθμος εκπαίδευσης. Η επιλογή του αλγορίθμου εκπαίδευσης καθορίζεται από τη δομή του κάθε δικτύου καθώς και από το

εκάστοτε πρόβλημα. Σύμφωνα με τον δεύτερο τρόπο η εκπαίδευση πραγματοποιείται χωρίς εποπτεία. Πιο συγκεκριμένα, κάθε δίκτυο που εκπαιδεύεται χωρίς εποπτεία χρειάζεται να αντιλαμβάνεται ομοιότητες σε πληροφορίες και δεδομένα με τα οποία έχει τροφοδοτηθεί. Το δίκτυο χρειάζεται να προσαρμοστεί και να χωρίσει τα δεδομένα σε ομάδες. Η διαδικασία αυτή συνεχίζεται μέχρις ότου σταματήσει η ταξινόμηση των δεδομένων.

Η λειτουργία κάθε τεχνητού νευρωνικού δικτύου εξαρτάται από διάφορους παράγοντες. Αναλυτικότερα, οι παράγοντες αυτοί είναι τα ειδικά χαρακτηριστικά των νευρώνων του κάθε δικτύου, η τοπολογία του, τα δεδομένα και ο τρόπος εκπαίδευσής του. Όλοι οι νευρώνες ενός δικτύου λειτουργούν ταυτόχρονα και εκτελούν έναν κοινό και σχετικά απλό υπολογισμό. Για το λόγο αυτό τα συγκεκριμένα δίκτυα αποτελούν παράδειγμα ενός μαζικά παράλληλου υπολογισμού.

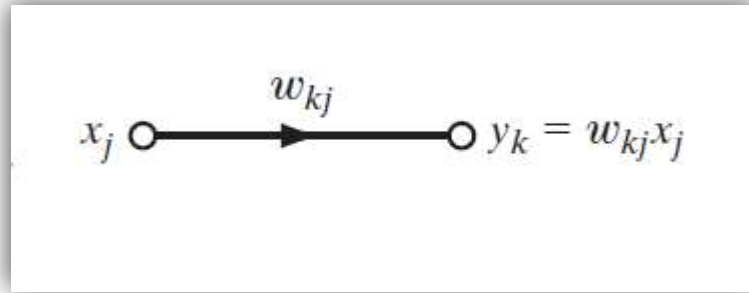
#### 1.4 Νευρωνικά Δίκτυα ως κατευθυνόμενοι γράφοι

Κάθε νευρωνικό δίκτυο μπορεί να θεωρηθεί και ως ένας κατευθυνόμενος γράφος. Πρόκειται για έναν γράφο, ο οποίος απαρτίζεται από κόμβους συναπτικών συνδέσεων και ενεργοποίησης. Πιο συγκεκριμένα, ένα νευρωνικό δίκτυο αναπαριστάνεται με ένα πλήθος γραμμικών συναπτικών διασυνδέσεων, μια μη-γραμμική σύνδεση για την ενεργοποίηση του καθώς και με ένα κατώφλι.

Το Νευρωνικό Δίκτυο ως κατευθυνόμενος γράφος διαθέτει πολλές και χρήσιμες ιδιότητες. Με σκοπό την απλοποίηση όσον αφορά την εμφάνιση ενός τεχνητού νευρώνα, γίνεται χρήση των γράφων ροής σημάτων. Κάθε σήμα μέσα στο Τεχνητό Νευρωνικό δίκτυο ρέει προς την μοναδική κατεύθυνση που υποδεικνύει το βέλος. (Haykin, 2009) Διακρίνονται δύο διαφορετικοί τύποι συνδέσεων, οι οποίοι είναι οι εξής:

- Συναπτικές Συνδέσεις

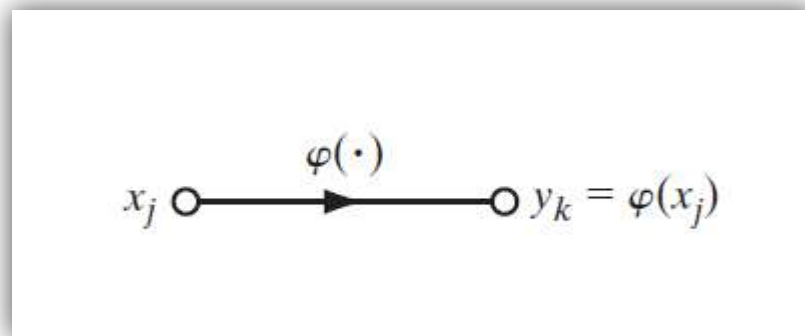
Στην περίπτωση αυτή, προκειμένου να παραχθεί το τελικό σήμα ή η έξοδος  $y_k$  ένα σήμα  $x_j$  πολλαπλασιάζεται με το βάρος του  $w_{kj}$ , όπως φαίνεται στην παρακάτω εικόνα.



**Εικόνα 1 - Συναπτική σύνδεση**

- Συνδέσεις ενεργοποίησης

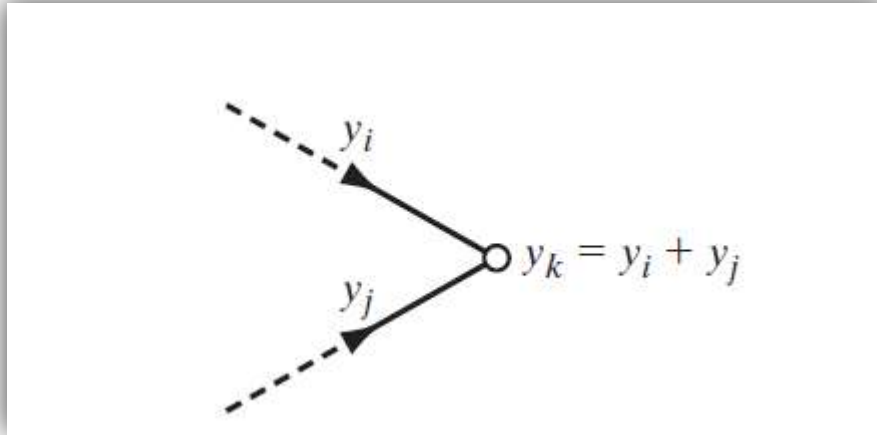
Στη συγκεκριμένη περίπτωση, με τις συνδέσεις ενεργοποίησης αναφερόμαστε σε μια μη γραμμική συνάρτηση, η οποία συμβολίζεται με  $\Phi(\cdot)$  όπως φαίνεται και από την Εικόνα 2 που ακολουθεί. (Haykin, 2009)



**Εικόνα 2 - Συνδέσεις ενεργοποίησης**

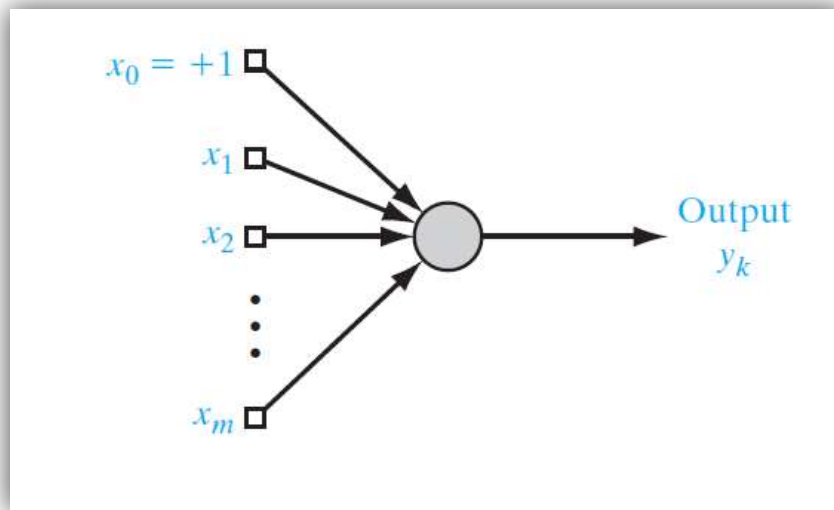
Μια ακόμα πολύ σημαντική ιδιότητα σε ένα νευρωνικό δίκτυο είναι ότι ένα σήμα κόμβου είναι ίσο με το άθροισμα των σημάτων που καταλήγουν στον συγκεκριμένο κόμβο όπως απεικονίζεται από την εικόνα 3 που βρίσκεται στη συνέχεια.





**Εικόνα 3 - Απεικόνιση σήματος κόμβου**

Στα Νευρωνικά δίκτυα υπάρχουν επίσης οι λεγόμενοι «αρχιτεκτονικοί γράφοι». Πρόκειται ουσιαστικά για μερικώς ολοκληρωμένους κατευθυνόμενους γράφους. Στην Εικόνα 4 που ακολουθεί αναπαριστάνεται ένας αρχιτεκτονικός γράφος.

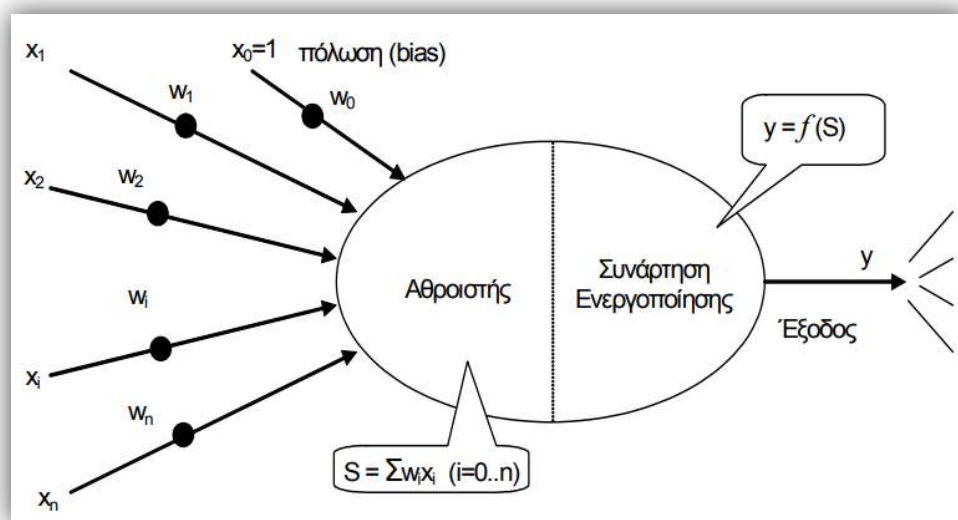


**Εικόνα 4 - Αρχιτεκτονικός Γράφος**

### 1.5 Μοντέλο τεχνητού νευρώνα

Οι νευρώνες αποτελούν το πιο βασικό μέρος ενός τεχνητού νευρωνικού δικτύου. Για το λόγο αυτό αξίζει να αναφερθούμε στους νευρώνες πιο αναλυτικά

και να αναλυθούν η δομή και η λειτουργία τους περαιτέρω. Με τον όρο τεχνητός νευρώνας, αναφερόμαστε σε μια μονάδα επεξεργασίας των πληροφοριών και των δεδομένων. Διαχειρίζονται εξ ολοκλήρου τοπικά δεδομένων. Τα δεδομένα αυτά εισέρχονται από τις συνδέσεις στους νευρώνες ως εισόδοι. Ο νευρώνας αποτελείται από τρία βασικά δομικά στοιχεία, τις συνάψεις, τον αθροιστή και τη συνάρτηση ενεργοποίησης, όπως φαίνεται και από την **Εικόνα 5** που ακολουθεί.



**Εικόνα 5 - Μοντέλο Τεχνητού Νευρώνα**

Πιο αναλυτικά, τα στοιχεία αυτά είναι :

**Συνάψεις:** Κάθε τεχνητός νευρώνας περιλαμβάνει ένα πλήθος από συνάψεις ή αλλιώς συνδετικούς κρίκους (P.J. Braspenning, A.J.M.M. Weijters, F. Thuijsman, 1995 ). Όλες οι συνάψεις του νευρώνα διαθέτουν ένα χαρακτηριστικό βάρος. Αν μια σύναψη είναι διεγερτική, τότε το βάρος του, που συμβολίζεται με  $w$  είναι θετικό, ενώ στην αντίθετη περίπτωση όπου μια σύναψη είναι απαγορευτική, το βάρος  $w$  είναι αρνητικό.

**Αθροιστής:** Ένα από τα βασικά στοιχεία ενός τεχνητού νευρώνα, όπως ήδη αναφέρθηκε είναι και ο αθροιστής. Σκοπός του αθροιστή είναι η πρόσθεση όλων των σημάτων εισόδου του νευρώνα, τα οποία έχουν πολλαπλασιαστεί με τα βάρη των συνδέσεων. Η συγκεκριμένη διεργασία αποτελεί μια γραμμική λειτουργία.

**Συνάρτηση ενεργοποίησης:** Στον τεχνητό νευρώνα υπάρχει μια συνάρτηση ενεργοποίησης, η οποία χρησιμεύει στη ελάττωση του εύρους στην έξοδο του κάθε νευρώνα. Συγκεκριμένα, ορίζεται η έξοδος σε συνάρτηση με τον βαθμό ενεργοποίησης της εισόδου. Κάθε νευρώνας, προκειμένου να κατασκευάσει κατάλληλα την έξοδο του, έχει τη δυνατότητα να χρησιμοποιήσει διαφορετική συνάρτηση. (Haykin, 2009)

Οι μαθηματικές σχέσεις που περιγράφουν έναν τεχνητό νευρώνα είναι οι ακόλουθες:

$$u_k = \sum_{j=0}^m w_{kj} x_j$$

και

$$y_k = \varphi(u_k)$$

Κάθε συνάρτηση ενεργοποίησης συμβολίζεται με  $\varphi(u)$ . Στη συνέχεια, αναφέρονται και περιγράφονται δύο βασικές συναρτήσεις ενεργοποίησης:

- **Συνάρτηση Κατωφλιού (Threshold Function)**

Η συνάρτηση ενεργοποίησης κατωφλιού είναι η ακόλουθη:

$$\varphi(u) = \begin{cases} 1 & \text{αν } u \geq 0 \\ 0 & \text{αν } u < 0 \end{cases}$$

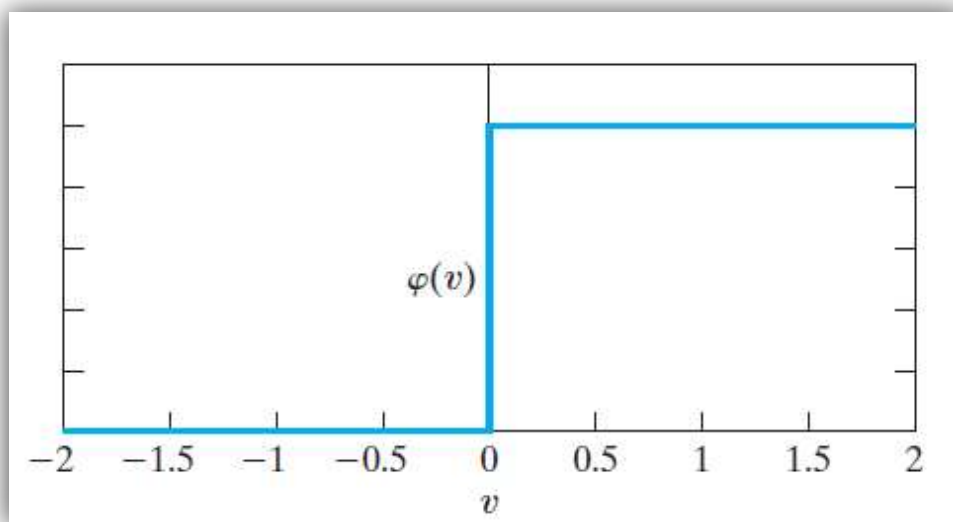
Για την έξοδο ενός νευρώνα έστω  $k$ , η συνάρτηση κατωφλιού αναφέρεται ως:

$$y_k = \begin{cases} 1 & \text{αν } u_k \geq 0 \\ 0 & \text{αν } u_k < 0 \end{cases}$$

όπου

$$u_k = \sum_{j=1}^m w_{kj} x_j + b_k$$

Στην Εικόνα 6, που βρίσκεται παρακάτω, παρουσιάζετε η γραφική παράσταση της συνάρτησης κατωφλιού. Στον κλάδο των νευρωνικών δικτύων, ένας νευρώνα που χρησιμοποιεί συνάρτηση ενεργοποίησης κατωφλιού αναφέρονται ως μοντέλο McCulloch-Pitts, από τις εργασίες των McCulloch και Pitts το 1943. Σύμφωνα με το συγκεκριμένο μοντέλο, η έξοδος κάθε νευρώνα δίνει ως αποτέλεσμα 1 εάν η υπολογισμένη τιμή του αθροιστή  $u_k$  είναι μη αρνητική. Σε αντίθετη περίπτωση η τιμή της εξόδου του νευρώνα παίρνει την τιμή 0. (Haykin, 2009)



**Εικόνα 6 - Συνάρτηση Ενεργοποίησης Κατωφλιού**

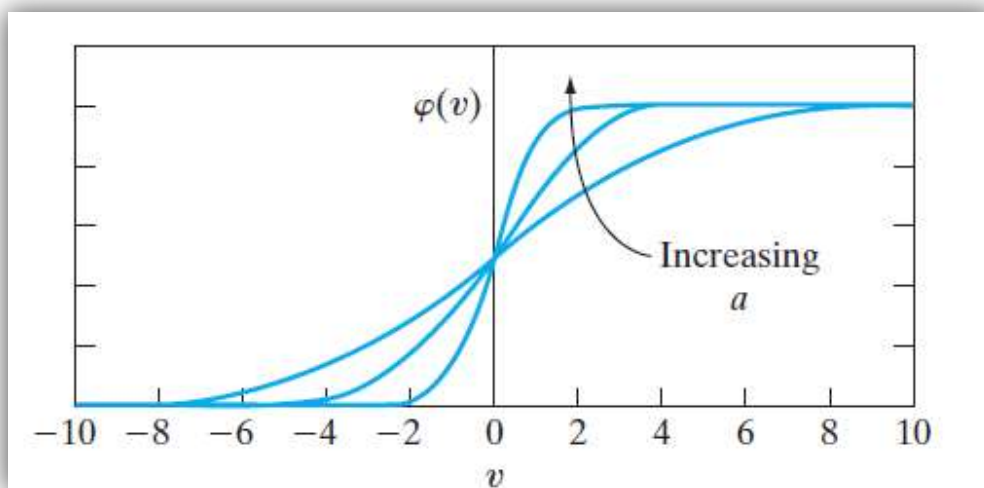
Στα πρώτα μοντέλα τεχνητών νευρώνων που αναπτύχθηκαν, χρησιμοποιήθηκε ως συνάρτηση ενεργοποίησης η βηματική συνάρτηση (step function), η οποία είναι μια συνάρτηση κατωφλιού. Κατά τη βηματική συνάρτηση, η έξοδος ενός νευρώνα είχε την τιμή 1 αν το αποτέλεσμα έδινε τιμή μεγαλύτερη από μια προκαθορισμένη τιμή κατωφλιού και ο νευρώνας ήταν ενεργοποιημένος. Διαφορετικά, εάν το υπολογισμένο αποτέλεσμα ήταν μικρότερο από την τιμή κατωφλιού, τότε η έξοδος είχε την τιμή 0 και ο νευρώνας ήταν αδρανής.

Με την πάροδο του χρόνου, οι επιστήμονες που ασχολήθηκαν με τον τομέα των νευρωνικών δικτύων, ανακάλυψαν την μεγάλη σημασία που είχε η παράγωγος της συνάρτησης ενεργοποίησης. Μέσα από την παράγωγό της

μπορούσαν να βγουν σημαντικές πληροφορίες για το νευρωνικό δίκτυο που θα ωφελούσαν στη διαδικασία μάθησης του δικτύου. Κατέληξαν στο συμπέρασμα ότι είναι πιο ωφέλιμο να χρησιμοποιούνται συναρτήσεις ενεργοποίησης, οι οποίες είναι παραγωγίσιμες, εν αντιθέσει με μη παραγωγίσιμες συναρτήσεις όπως είναι η βηματική συνάρτηση. (Αργυράκης, 2001)

- **Σιγμοειδής Συνάρτηση (Sigmoid Function)**

Στις μέρες μας, στα περισσότερα τεχνητά νευρωνικά δίκτυα χρησιμοποιείτε ως συνάρτηση ενεργοποίησης μια σιγμοειδή συνάρτηση. Πρόκειται για μια συνεχής, πραγματική συνάρτηση. Η παράγωγος της συγκεκριμένης συνάρτησης είναι θετική. Στην Εικόνα 7, που βρίσκεται παρακάτω, παρουσιάζετε η γραφική παράσταση της σιγμοειδής συνάρτησης ενεργοποίησης. Όπως φαίνεται και από την εικόνα η συγκεκριμένη συνάρτηση έχει τη μορφή γραφικά ενός “S”.



**Εικόνα 7 - Σιγμοειδής Συνάρτηση Ενεργοποίησης**

Ένα παράδειγμα σιγμοειδής συνάρτησης ενεργοποίησης μπορούμε να αναφέρομαι τη λογιστική συνάρτηση (logistic function). Ο τύπος της είναι ο ακόλουθος:

$$\varphi(u) = \frac{1}{1 + \exp(-au)}$$

Στον παραπάνω τύπο με την μεταβλητή  $\alpha$  αναφερόμαστε στην παράμετρο κλίσης. Αλλάζοντας την τιμή της παραμέτρου  $\alpha$ , δημιουργούνται συναρτήσεις διαφορετικών κλίσεων. Η λογιστική συνάρτηση τείνει στην βηματική, όταν η παράμετρος  $\alpha$  τείνει προς το άπειρο. (Αργυράκης, 2001)

Ως επιπλέον παραδείγματα σιγμοειδών συναρτήσεων μπορούν να αναφερθούν η συνάρτηση τόξου εφαπτομένης (arctangent function) και η υπερβολική συνάρτηση (hyperbolic function).

Οι συναρτήσεις που περιγράφηκαν προηγουμένως αφορούν τις τιμές που βρίσκονται από 0 έως +1. Εάν απαιτείται μια συνάρτηση ενεργοποίησης, στην οποία οι τιμές πρέπει να κυμαίνονται από -1 έως +1 επιλέγετε ως συνάρτηση κατωφλιού η συνάρτηση προσήμου. Η συνάρτηση προσήμου συμβολίζεται ως  $\text{sgn}()$  και ο τύπος της είναι ο εξής:

$$\varphi(u) = \begin{cases} 1, & u > 0 \\ 0, & u = 0 \\ -1, & u < 0 \end{cases}$$

Οι συναρτήσεις που χρησιμοποιούν οι νευρώνες ενός τεχνητού νευρωνικού δικτύου είναι μη γραμμικές συναρτήσεις, οι οποίες συμβάλλουν στη δημιουργία της εξόδου. Η σύνδεση των διαφορετικών επιπέδων ενός δικτύου καθορίζεται από τα συναπτικά βάρη. Τα συγκεκριμένα βάρη που αναφέρονται έχουν προσαρμοστεί σύμφωνα με έναν αλγόριθμο μάθησης, προκειμένου να εκπαιδευτεί σωστά και κατάλληλα ένα Νευρωνικό Δίκτυο. (Haykin, 2009)

## 1.6 Ικανότητα γενίκευσης

Μια από τις πιο σημαντικές ιδιότητες που διακρίνεται στα Τεχνητά Νευρωνικά Δίκτυα είναι η ιδιότητα της γενίκευσης. Με τον όρο γενίκευση αναφερόμαστε στην σημαντική ικανότητα του δικτύου να μπορεί να αποδώσει σωστά για πρότυπα εισόδου που δεν έχουν χρησιμοποιηθεί ως πρότυπα εκπαίδευσης. Πρόκειται για μια σημαντική ικανότητα, διότι σχεδόν για όλα τα προβλήματα, τα οποία καλούνται τα νευρωνικά δίκτυα να επιλύσουν δεν είναι γνωστές από την αρχή οι καταστάσεις που θα προκύψουν. (N Ζώρης, 2014)

Ο όρος της γενίκευσης που ήδη αναφέρθηκε είναι ένας όρος που προέρχεται και δανείζεται από την ψυχολογία. Διακρίνεται μια σημαντική διαφορά μεταξύ του τμήματος της επεξεργασίας της πληροφορίας (pattern classifier) με την υλοποίηση ενός Τεχνητού Νευρωνικού Δικτύου.

Για την πρώτη περίπτωση για την επεξεργασία της πληροφορίας η διαδικασία ξεκινάει με τον σχηματισμό ενός μαθηματικού προτύπου περιβαλλοντικών κριτηρίων. Στη συνέχεια επαληθεύεται με πραγματικά δεδομένα και τέλος με βάση το συγκεκριμένο μοντέλο υλοποιείται ο σχεδιασμός του. Για τη δεύτερη περίπτωση του σχεδιασμού ενός Νευρωνικού Δικτύου η διαδικασία βασίζεται σε πραγματικά δεδομένα, τα οποία συμβάλλουν στην όλη διαδικασία.

Σύμφωνα με όλα που προαναφέρθηκαν, φτάνουμε στο συμπέρασμα ότι το νευρωνικό δίκτυο παρέχει ένα αξιόπιστο μοντέλο στο περιβάλλον που λειτουργεί καθώς και δίνει τη δυνατότητα επεξεργασίας των πληροφοριών.

Η σημαντική ιδιότητα της γενίκευσης μπορεί να επηρεαστεί για διάφορους λόγους. Στη συνέχεια, αναφέρονται ορισμένοι παράγοντες που πιθανώς να την επηρεάσουν:

- Διανύσματα μάθησης.

Η γενίκευση μπορεί να επηρεαστεί τόσο από το πλήθος των διανυσμάτων που υπάρχουν για τη διαδικασία της μάθησης όσο και από το αν η αναπαράσταση του περιβάλλοντός τους είναι ικανοποιητική. Η καλύτερη λύση είναι να υπάρχει ένα μεγάλο πλήθος των διανυσμάτων μάθησης. Με τον τρόπο αυτό μπορεί να επιτευχθεί η καλύτερη λύση.

- Μέγεθος του Νευρωνικού Δικτύου που χρησιμοποιείται.
- Πολυπλοκότητα του εκάστοτε προβλήματος που καλείται να υλοποιήσει το Νευρωνικό Δίκτυο.

## **1.7 Βιολογικά Νευρωνικά Δίκτυα**

Η ιδέα για τη δημιουργία και την κατασκευή των νευρωνικών δικτύων βασίστηκε στις ικανότητες και τις λειτουργίες του εγκεφάλου των ζωντανών οργανισμών. Για το λόγο αυτό, είναι ιδιαίτερα χρήσιμο και απαραίτητο να αναφερθούμε στη δομή και τη λειτουργία του.

### 1.7.1 Εγκέφαλος: ένα σύνθετο βιολογικό νευρωνικό δίκτυο

Τα πρώτα δίκτυα που μελετήθηκαν από τους επιστήμονες ήταν τα νευρωνικά δίκτυα των ζωντανών οργανισμών, τα οποία αναφέρονται ως Βιολογικά Νευρωνικά Δίκτυα. Πρωταρχικός στόχος για τη δημιουργία των τεχνητών νευρωνικών δικτύων ήταν η δημιουργία ενός μοντέλου που θα προσομοίωνε όσο το δυνατόν καλύτερα την πολύπλοκη λειτουργία του ανθρώπινου εγκεφάλου. (P.J. Braspenning, A.J.M.M. Weijters, F. Thuijisman, 1995 )

Όπως έχει ήδη αναφερθεί, η ιδέα για κάθε πρότυπο νευρωνικού δικτύου ξεκίνησε και βασίστηκε στη βιολογία. Όλοι οι ζωντανοί οργανισμοί, όπως τα ζώα και οι άνθρωποι διαθέτουν ένα νευρικό σύστημα το οποίο εκτελεί διάφορες διεργασίες. Ως παραδείγματα τέτοιων διεργασιών μπορούμε να αναφέρουμε τη διαδικασία μάθησης, τη δυνατότητα απομνημόνευσης καθώς και την επαφή με το εξωτερικό περιβάλλον. Κάθε νευρικό σύστημα των οργανισμών διαθέτει ένα μεγάλο αριθμό νευρωνικών δικτύων, τα οποία είναι υπεύθυνα να εκτελούν συγκεκριμένες λειτουργίες. Ο εγκέφαλος αποτελεί την κεντρική μονάδα αυτού του συστήματος και απαρτίζεται επίσης από ένα μεγάλο αριθμό νευρωνικών δικτύων. (Haykin, 2009)

Ο ανθρώπινος εγκέφαλος αποτελεί ένα σύνθετο, βιολογικό νευρωνικό δίκτυο. Η δομή και η λειτουργία του είναι τέτοια, ώστε να επιτρέπει την σύνθετη επεξεργασία πολλών δεδομένων και πληροφοριών ταυτόχρονα. (Rojas, 1996) Ένα δεύτερο σημαντικό χαρακτηριστικό είναι η συνεχή μάθηση αλληλεπιδρώντας συνεχώς με το εξωτερικό περιβάλλον. Ανταποκρίνεται διαρκώς σε καινούργιες εμπειρίες, βασιζόμενος σε αντίστοιχες, προηγούμενες αποθηκευμένες εμπειρίες. Ο εγκέφαλος των ανθρώπων κατέχει την ικανότητα της γενίκευσης, δηλαδή εκτελεί προβλέψεις για καινούργιες καταστάσεις (Haykin, 2009). Προκειμένου να ληφθεί μια απόφαση δεν χρειάζονται πλήρης δεδομένα και πληροφορίες. Ο λόγος που συμβαίνει αυτό είναι ότι ο εγκέφαλος έχει αντοχές όσον αφορά τα ανεπαρκή δεδομένα καθώς και την παραμόρφωση αυτών. Διαθέτει μεγάλη ανοχή στις βλάβες. Πιο συγκεκριμένα, η πιθανή απώλεια ορισμένων νευρώνων αντιμετωπίζεται με τη σωστή διαμόρφωση των υπόλοιπων νευρώνων του δικτύου, σε συνδυασμό με την επιπλέον εκπαίδευση. Με αυτά τα χαρακτηριστικά ο ανθρώπινος εγκέφαλος εκτελεί δύσκολες λειτουργίες και συνεχώς εξελίσσεται.

Το συνολικό πλήθος των νευρικών κυττάρων σε έναν ζωντανό οργανισμό είναι πολύ μεγάλο και δεν μπορεί να προσδιοριστεί με ακρίβεια. Οι επιστήμονες όμως υποστήριξαν ότι ο εγκέφαλος απαρτίζεται από περίπου δυο τρισεκατομμύρια νευρώνες. (Rojas, 1996). Κατά τη γέννηση ενός ζωντανού οργανισμού, ο εγκέφαλος δημιουργεί απευθείας τους δικούς του κανόνες, οι οποίοι διαμορφώνονται και εξελίσσονται με την πάροδο του χρόνου. Η



μεγαλύτερη ανάπτυξη παρατηρείτε τα δύο πρώτα χρόνια ζωής, στα οποία ο κάθε οργανισμός δημιουργεί ένα εκατομμύριο περίπου συνάψεις (synapses) που μεσολαβούν στην επικοινωνία των νευρώνων.

Διακρίνονται δύο βασικές ομοιότητες των Νευρωνικών δικτύων με τον ανθρώπινο εγκέφαλο:

- Οι γνώσεις και οι εμπειρίες που αποκτούν τόσο τα Νευρωνικά δίκτυα όσο και ο εγκέφαλος των ζωντανών οργανισμών λαμβάνονται από το περιβάλλον τους αντίστοιχα μέσα από μια διαδικασία εκπαίδευσης τους. (Haykin, 2009)
- Τόσο στα Νευρωνικά δίκτυα όσο και στον ανθρώπινο εγκέφαλο υπάρχουν οι αντίστοιχες συνδέσεις. Και στις δύο περιπτώσεις οι συνδέσεις αυτές που είναι τα συναπτικά βάρη χρησιμεύουν για την αποθήκευση της γνώσης που αποκτάτε κάθε φορά. (Haykin, 2009)

### 1.7.2 Βιολογικός Νευρώνας

Όλα τα βιολογικά νευρωνικά δίκτυα αποτελούνται από μια πληθώρα μονάδων, οι οποίες ονομάστηκαν νευρώνες (neurons). Στην Εικόνα 8 απεικονίζεται η δομή ενός βιολογικού νευρώνα. Η κυριότερη λειτουργία τους είναι η συνεχής επεξεργασία των πληροφοριών, ενώ δέχονται και στέλνουν διαρκώς σήματα από και προς άλλους νευρώνες. Όπως παρατηρείτε και από την Εικόνα 8, τα κύρια δομικά στοιχεία κάθε βιολογικού νευρώνα είναι το Σώμα, οι Δενδρίτες καθώς και ο Νευράξονα. (Πλέρου, 2012)

Πιο αναλυτικά:

- Το **Σώμα** αποτελεί το μεγαλύτερο τμήμα του νευρώνα συγκριτικά με τον Νευράξονα και τους Δενδρίτες. Ο πυρήνας του κυττάρου περιλαμβάνεται επίσης στο Σώμα.
- Οι **Δενδρίτες** είναι ουσιαστικά μικρές, διακλαδιζόμενες κυτταρικές προεξοχές.

- Ο **Νευράξονας** αναφέρεται επίσης και ως Νευρίτης. Πρόκειται για μια λεπτή ίνα, η οποία έχει τη δυνατότητα να μεταφέρει τα σήματα από τον νευρώνα. Υπάρχει μόνο ένας Νευράξονας σε κάθε νευρώνα. Στο τέλος κάθε άξονα υπάρχουν σημαντικοί κόμβοι, από τους οποίους μεταφέρονται οι πληροφορίες σε άλλους δενδρίτες επόμενων νευρώνων.

Σε κάθε βιολογικό νευρώνα διακρίνονται δύο καταστάσεις. Η πρώτη κατάσταση είναι η **ενεργή**, κατά την οποία ο νευρώνας συνεχώς πυροδοτεί ηλεκτρικούς παλμούς. Οι παλμοί αυτοί διαδίδονται μέσω των συνάψεων και διεγείρουν δενδρίτες άλλων νευρώνων. Όταν το πλήθος των διεγέρσεων αυτών σε έναν νευρώνα φτάσει ένα συγκεκριμένο κατώφλι, τότε και μόνο αυτός πυροδοτεί. Η δεύτερη κατάσταση ενός νευρώνα είναι η κατάσταση της **αδράνειας**. (Wilde, 1996)



**Εικόνα 8 - Βιολογικός Νευρώνας**

Κάθε νευρώνας απαρτίζεται από ένα μεγάλο αριθμό δενδριτών, από τα οποία το καθένα αποτελείται από πολλές διακλαδώσεις. Με βάση αυτά τα βασικά δομικά στοιχεία του νευρώνα που αναφέρθηκαν, κάθε νευρώνας δέχεται τα σήματα από άλλα κύτταρα. Πιο συγκεκριμένα, σήματα φτάνουν συνεχώς σε έναν δενδρίτη από έναν ή και περισσότερους νευρίτες άλλων κυττάρων. Επομένως, συνδέονται και συνεργάζονται μεταξύ τους με σκοπό την ομαλή λειτουργία τους για τη διάδοση των ερεθισμάτων από και προς τον εγκέφαλο. Τα προσυναπτικά κομβία καθώς και οι συνάψεις βρίσκονται στις άκρες των δενδριτών, προκειμένου να δέχονται ή να στέλνουν τα επιθυμητά σήματα. (Haykin, 2009)

### 1.7.3 Υπολογιστής και Ανθρώπινος Εγκέφαλος

Συγκρίνοντας ένα υπολογιστή με τον ανθρώπινο εγκέφαλο έχει παρατηρηθεί διαφορά στην κατανάλωση ενέργειας. Πιο συγκεκριμένα, η ενέργεια που καταναλώνεται στον εγκέφαλο είναι 20 Watt, ενώ αντίθετα ένας υπολογιστής χρειάζεται πολύ περισσότερη ενέργεια της τάξης των χιλιάδων Watt. (Πλέρου, 2012)

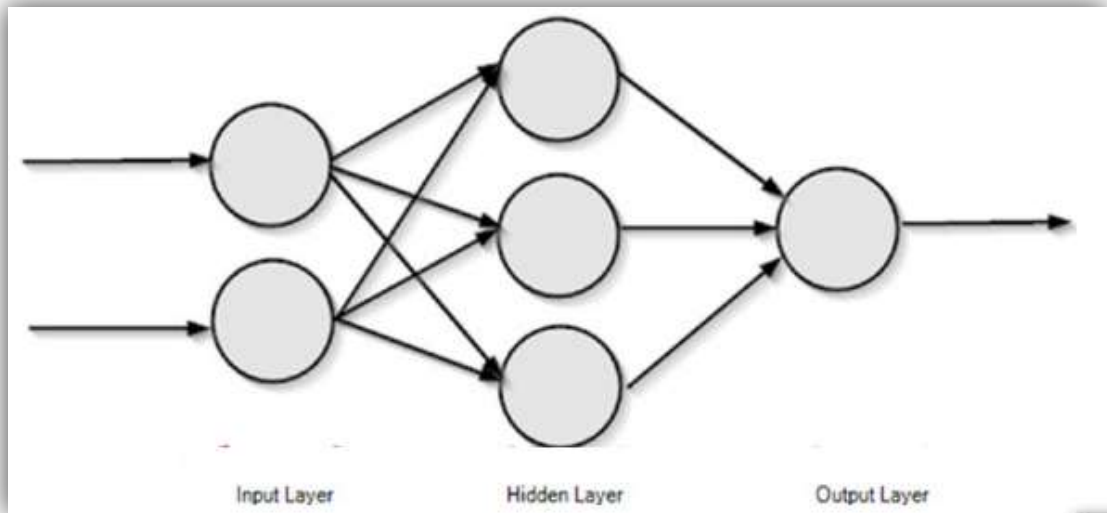
Έκτος από ορισμένες διαφορές που υπάρχουν ανάμεσα σε ένα τεχνητό νευρωνικό δίκτυο και τον ανθρώπινο εγκέφαλο υπάρχουν και πολλές ομοιότητες. Μια πολύ σημαντική ομοιότητα παρατηρείτε στον τρόπο μάθησης. Κάθε δίκτυο είτε τεχνητό είτε βιολογικό αποκτά γνώσεις σύμφωνα με μια διαδικασία μάθησης. Η βασική τους διαφορά ως προς τον τρόπο μάθησης όμως είναι ότι τα τεχνητά νευρωνικά δίκτυα εν αντιθέσει με τα βιολογικά δίκτυα βασίζονται σε συγκεκριμένους προκαθορισμένους κανόνες.

## 1.8 Αρχιτεκτονικές νευρωνικών δικτύων

Οι νευρώνες στα τεχνητά νευρωνικά δίκτυα συνδέονται μεταξύ τους μέσω διαφόρων αρχιτεκτονικών. Η επιλογή της αρχιτεκτονικής καθορίζεται από τη διαδικασία μετάδοσης των πληροφοριών ανάμεσα στους διάφορους νευρώνες. Στη συνέχεια αναλύονται οι βασικές αρχιτεκτονικές των νευρωνικών δικτύων.

### 1.8.1 Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης (feedforward)

Στα Τεχνητά Νευρωνικά δίκτυα πρόσθιας τροφοδότησης, τα σήματα διαδίδονται αποκλειστικά και μόνο με κατεύθυνση από την είσοδο του νευρωνικού δικτύου προς την έξοδο. (Haykin, 2009) Στη συγκεκριμένη αρχιτεκτονική δεν υπάρχει σε κανένα σημείο ανατροφοδότηση. Δεν υπάρχει μέσα στο Νευρωνικό δίκτυο κανένας νευρώνας, του οποίου η έξοδος αποτελεί είσοδο άλλου νευρώνα στο ίδιο ή στο προηγούμενο επίπεδο. Στην Εικόνα 9 αναπαριστάται ένα πρότυπο δικτύου πρόσθιας τροφοδότησης, στο οποίο δεν υπάρχουν συνδέσεις με κλειστά μονοπάτια.

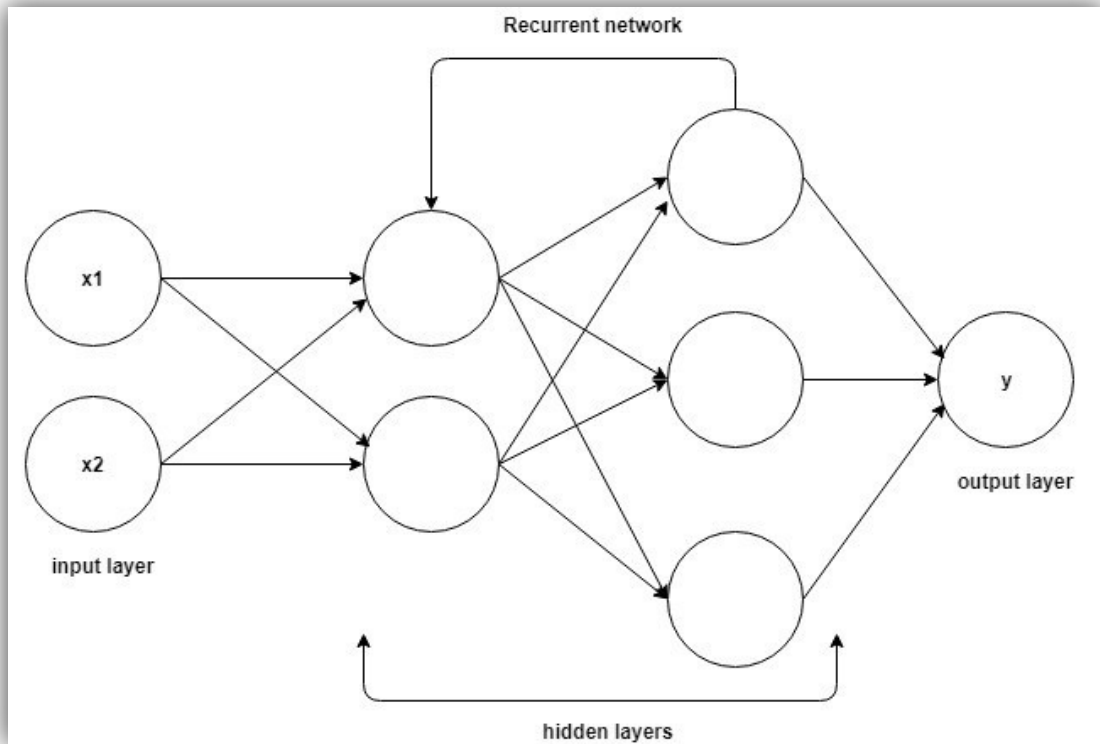


**Εικόνα 9 - Τεχνητό Νευρωνικό Δίκτυο πρόσθιας τροφοδότησης**

### **1.8.2 Νευρωνικά Δίκτυα με ανατροφοδότηση (feedback, recurrent)**

Στα Τεχνητά Νευρωνικά Δίκτυα με ανατροφοδότηση, τα σήματα διαδίδονται και στις δύο κατευθύνσεις, μέσω των βρόγχων που υπάρχουν στο δίκτυο. (Haykin, 2009) Από τα βασικά χαρακτηριστικά των συγκεκριμένων δικτύων είναι η λειτουργικότητα τους καθώς και το γεγονός ότι είναι πολύ ισχυρά. Από την άλλη τα δίκτυα αυτά θεωρούνται ιδιαίτερα πολύπλοκα. Η κατάσταση τους μεταβάλλεται διαρκώς μέχρις ότου βρεθούν σε μια κατάσταση ισορροπίας. Παραμένουν στο σημείο ισορροπίας ώσπου να τροποποιηθεί η έξοδος. Στο σημείο αυτό θα βρεθεί καινούργιο σημείο ισορροπίας.

Τα δίκτυα με τη συγκεκριμένη αρχιτεκτονική διαθέτουν πολλές δυνατότητες. Ένα μειονέκτημα που διακρίνεται στα δίκτυα με αρχιτεκτονική ανατροφοδότησης είναι ότι αντιμετωπίζονται μαθηματικά με περισσότερες δυσκολίες. Στην Εικόνα 10 αναπαριστάται ένα πρότυπο δικτύου ανατροφοδότησης, στο οποίο επιτρέπεται οι συνδέσεις των νευρώνων να σχηματίζουν κλειστό κύκλο.



**Εικόνα 10 - Τεχνητό Νευρωνικό Δίκτυο Ανατροφοδότησης**

### 1.9 Πλεονεκτήματα

Τα Νευρωνικά Δίκτυα παρουσιάζονται πολλά πλεονεκτήματα σύμφωνα με τις ιδιότητες και τις λειτουργίες τους. Στη συνέχεια αναφέρονται τα πιο σημαντικά πλεονεκτήματα τους.

#### **Μη-γραμμικότητα**

Μια πολύ σημαντική ιδιότητα των νευρωνικών δικτύων είναι η μη γραμμικότητα. Πιο συγκεκριμένα, κάθε νευρωνικό δίκτυο είναι δομημένο με συνδεδεμένους νευρώνες, οι οποίοι αποτελούν συσκευές μη-γραμμικές. (Haykin, 2009)

#### **Απομνημόνευση πληροφοριών**

Ένα από τα πιο σημαντικά πλεονεκτήματα που παρατηρείτε στα νευρωνικά δίκτυα είναι η δυνατότητα αποθήκευσης της εμπειρίας και της γνώσης από το περιβάλλον τους. Τις αποθηκευμένες αυτές πληροφορίες μπορεί να τις ανακαλέσει. Επίσης πολύ σημαντικό είναι ότι τα συγκεκριμένα δίκτυα έχουν τη δυνατότητα να εξάγουν ορισμένα βασικά στοιχεία ενός συστήματος, ακόμα και αν αυτά βρίσκονται σε θορυβώδη δεδομένα.

## **Προσαρμοστικότητα**

Ως πλεονέκτημα των νευρωνικών δικτύων μπορεί να αναφερθεί η ικανότητα της προσαρμοστικότητας. Όλα τα δίκτυα έχουν την δυνατότητα να εφαρμόζουν αλλαγές στα βάρη τους όταν παρατηρούνται μεταβολές στο περιβάλλον τους. Ένα δίκτυο, το οποίο ήταν εκπαιδευμένο να εκτελεί λειτουργίες σε ένα συγκεκριμένο περιβάλλον, μπορεί με μεγάλη ευκολία να εκπαιδευτεί εκ νέου προκειμένου να υλοποιήσει μικρές αλλαγές στις λειτουργίες του. (Haykin, 2009) Επιπλέον, ένα νευρωνικό δίκτυο μπορεί να σχεδιαστεί εξ αρχής με σκοπό να μπορεί να μεταβάλλει σε πραγματικό χρόνο τα συναπτόμενα βάρη του.

## **Αποδεικτική Ανταπόκριση**

Στα Νευρωνικά δίκτυα πραγματοποιείτε ταξινόμηση μεταξύ διαφόρων δειγμάτων. Τα δίκτυα αυτά μπορούν να σχεδιαστούν με τέτοιο τρόπο ώστε να δίνουν πληροφορίες όχι αποκλειστικά και μόνο για την επιλογή του δείγματος αλλά και κατά πόσο το συγκεκριμένο δείγμα είναι η κατάλληλότερη επιλογή. (Αργυράκης, 2001) Με τον τρόπο αυτό βελτιώνεται σημαντικά η απόδοση των δικτύων, καθώς με τις πληροφορίες αυτές θα μπορούσαν να βοηθήσουν στην απόρριψη ακατάλληλων δειγμάτων.

## **Ανοχή στα λάθη**

Στα συγκεκριμένα δίκτυα παρατηρείτε μεγάλη ανοχή στα λάθη. Συγκεκριμένα, κάνοντας σωστούς και κατάλληλους υπολογισμούς, πραγματοποιούνται λειτουργίες χωρίς μεγάλα σφάλματα.

## **Υλοποίηση με τεχνολογία VLSI**

Οι εφαρμογές που χρησιμοποιούν την τεχνολογία VLSI είναι οι πλέον κατάλληλες για τη φύση των νευρωνικών δικτύων. (Haykin, 2009)

### **1.10 Σύγχρονες Εφαρμογές**

Τα τελευταία χρόνια έχουν προκύψει αρκετές εφαρμογές των νευρωνικών δικτύων σε διάφορους τομείς. Ο αριθμός των εφαρμογών αυτών, τα επόμενα χρόνια αναμένεται να αυξηθεί σημαντικά, διότι ο τομέας των νευρωνικών δικτύων βρίσκονται σε πρώιμο στάδιο και εξελίσσεται συνεχώς. Στη συνέχεια περιγράφονται ορισμένες αντιπροσωπευτικές περιοχές εφαρμογής των νευρωνικών δικτύων. (N Ζώρης, 2014)

#### **Ιατρική**

Τα νευρωνικά δίκτυα έχουν βοηθήσει σημαντικά στον κλάδο της ιατρικής όσον αφορά την διαδικασία της διάγνωσης καθώς και τη θεραπεία. Πιο συγκεκριμένα, βοήθησε στην παρακολούθηση των εγχειρήσεων και στην ανάλυση της ομιλίας

με σκοπό την δημιουργία βοηθημάτων για ακουστικά προβλήματα. Συνέβαλλε, επίσης στην κατάλληλη φαρμακευτική πρόταση, βασιζόμενη στα συμπτώματα του κάθε ανθρώπου. Βοήθησε σημαντικά στην σωστή ανάγνωση των ακτινογραφιών. Τα νευρωνικά δίκτυα έδωσαν τη δυνατότητα για τη σωστή πρόβλεψη στη λήψη φαρμάκων όσον αφορά τις αντιδράσεις των ασθενών. (N Ζώρης, 2014)

### **Περιβάλλον**

Τα νευρωνικά δίκτυα έχουν βοηθήσει τους επιστήμονες όσον αφορά το περιβάλλον συμβάλλοντας στην ανάλυση των καιρικών συνθηκών και στην συνεχή πρόγνωση του καιρού.

### **Οικονομία**

Σημαντική συνεισφορά των νευρωνικών δικτύων παρατηρήθηκε και στον κλάδο της οικονομίας. Συγκεκριμένα ενίσχυσε τις διαδικασίες πρόβλεψης των τιμών στις μετοχές καθώς και της ανάλυσης της επικινδυνότητας των δανείων. Βοήθησε επίσης στη διαδικασία ανίχνευσης και καταπολέμησης κάθε είδους απάτης που συμβαίνει στις ηλεκτρονικές συναλλαγές.

### **Βιομηχανία**

Όσον αφορά τον κλάδο της Βιομηχανίας, τα δίκτυα συνέβαλλαν στη διαδικασία ελέγχου της παραγωγής, των χημικών διεργασιών και στον έλεγχο της ποιότητας. Δημιουργήθηκαν αυτοματοποιημένα ρομπότ και συστήματα, προκειμένου να πραγματοποιούνται οι κατάλληλοι έλεγχοι.

### **Σήματα**

Ο κλάδος αυτός συνέβαλε επίσης στην τεχνολογία των σημάτων και στην επεξεργασία τους. Πιο συγκεκριμένα, πραγματοποιήθηκε η μοντελοποίηση των σημάτων και ενισχύθηκε η διαδικασία απομάκρυνσης των θορύβων από τις τηλεφωνικές συνδέσεις.

### **Βιολογία**

Τα νευρωνικά δίκτυα βοήθησαν σημαντικά στη διαδικασία της εκμάθησης και κατανόησης των διαφόρων λειτουργιών του εγκεφάλου των ζωντανών οργανισμών. Ο κλάδος των νευρωνικών δικτύων έπαιξε, επίσης πολύ σημαντικό ρόλο στον τομέα της Βιολογία όσον αφορά την όραση και τη δημιουργία μοντέλων για τη βελτίωσή της.

### **Επιστήμη των Υπολογιστών**

Στον κλάδο της επιστήμης των Υπολογιστών, τα νευρωνικά δίκτυα συνέβαλλαν ως προς την αναγνώριση των προτύπων. Συγκεκριμένα αναφερόμαστε στην ικανότητα αναγνώρισης της φωνής, του προσώπου καθώς και των χειρόγραφων χαρακτήρων.

## Αεροπλοΐα

Τα νευρωνικά δίκτυα βοήθησαν επίσης σε μεγάλο βαθμό τον κλάδο της αεροπλοΐας. Δημιουργήθηκαν προγράμματα αυτόματου πιλότου καθώς και προγράμματα για την προσομοίωση μιας πτήσης. Κατασκευάστηκαν συστήματα με σκοπό τον έλεγχο μιας πτήσης και την ανίχνευση ελαττωμάτων και βλαβών σε οποιοδήποτε τμήμα των αεροσκαφών. (N Ζώρης, 2014).

### 1.11 Βασικές τεχνικές και εφαρμογές

Στην συγκεκριμένη υποενότητα θα αναφερθούν και θα αναλυθούν περαιτέρω ορισμένες εφαρμογές και παραδείγματα των Νευρωνικών Δικτύων σε επιστημονικούς κλάδους.

#### 1.11.1 Interpolation

Όπως έχει ήδη αναφερθεί, ένα νευρωνικό δίκτυο μπορεί χρησιμοποιηθεί για πολλούς λόγους. Ο πλέον απλός τρόπος χρήσης του είναι μια πολυμεταβλητής μη γραμμικής παλινδρόμησης. Η συγκεκριμένη μορφή ενός νευρωνικού δικτύου έχει ως σκοπό την εύρεση μιας ομαλής παρεμβολής αλληλεπίδρασης από ένα σύνολο δεδομένων.

Ως παράδειγμα, μπορεί να αναφερθεί ένα πρόβλημα για την πρόβλεψη μιας ποσότητας γνωστής ως χρόνου δέσμευσης ενέργειας  $\hat{t}_E$  ενός πλάσματος tokamak από τη γνώση των διαστάσεων τεσσάρων μεταβλητών. Αυτές οι τέσσερις μεταβλητές είναι οι ακόλουθες:

- $q$  - συντελεστής ασφάλειας.
- $v^*$ - αναλογία αποτελεσματικών συγκρούσεων ηλεκτρονίων προς συχνότητα ηλεκτρονίων.
- $\beta_p$ - ημιαγωγική βήτα. Αναλογία της πίεσης πλάσματος προς τη μαγνητική πίεση.
- $\hat{\rho}_e$ - κανονικοποιημένη ακτίνα Larmor ηλεκτρονίων.



Η συγκεκριμένη εφαρμογή που εξετάζεται σε αυτή την υποενότητα, στοχεύει στην πρόβλεψη του  $\hat{\tau}_E$  από τη γνώση των παραπάνω τεσσάρων μεταβλητών. Η σχέση που υπάρχει ανάμεσα σε αυτές τις μεταβλητές και το  $\hat{\tau}_E$  είναι η ακόλουθη.

$$\hat{\tau}_E = F (q, v^*, \beta_p, \hat{\rho}_e)$$

Στο σημείο αυτό, γίνεται μια αυθαίρετη υπόθεση ότι η συνάρτηση  $F ()$  που υπάρχει στην παραπάνω σχέση μπορεί να πάρει τη μορφή:

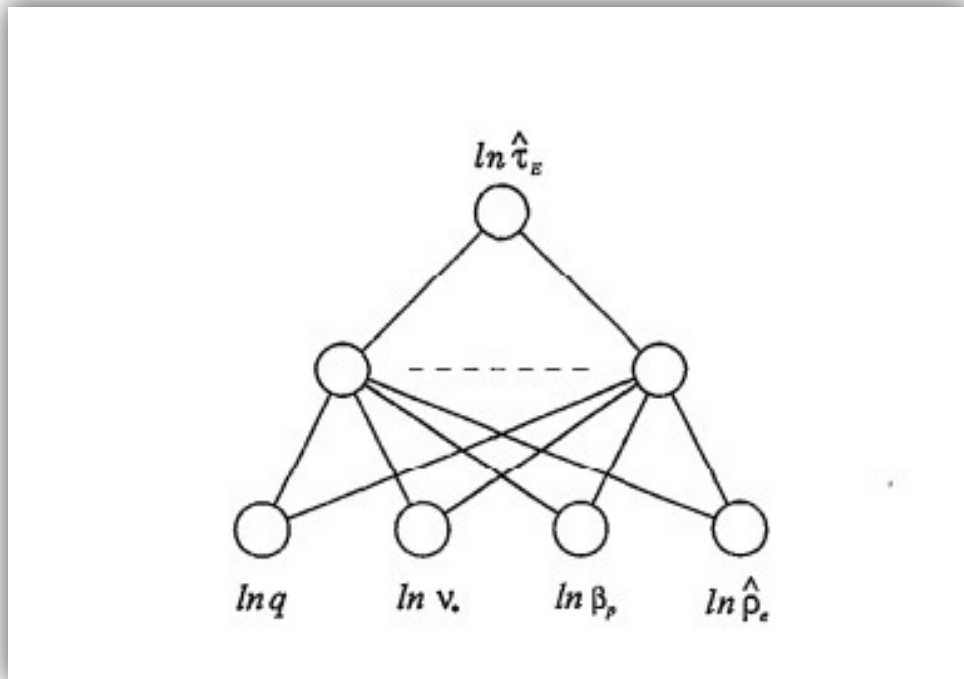
$$\hat{\tau}_E = e^C q^{a_1} v^{*a_2} \beta_p^{a_3} \hat{\rho}_e^{a_4}$$

Οι τιμές τόσο της παραμέτρου  $C$  όσο και των παραμέτρων  $a_1, a_2, a_3, a_4$  πρέπει να προσδιοριστούν από μια βάση δεδομένων. Εάν χρησιμοποιηθεί λογάριθμος στην παραπάνω σχέση, λαμβάνουμε μια έκφραση η οποία είναι γραμμική στις άγνωστες παραμέτρους.

$$\hat{\tau}_E = C + a_1 \ln q + a_2 \ln v^* + a_3 \ln \beta_p + a_4 \ln \hat{\rho}_e$$

Με τον τρόπο αυτό μπορούν να προσδιοριστούν οι παράμετροι από ένα σύνολο δεδομένων των τιμών ( $q, v^*, \beta_p, \hat{\rho}_e$ ), μαζί με τις αντίστοιχες τιμές του  $\hat{\tau}_E$ , με τις τεχνικές γραμμικής παλινδρόμησης.

Μπορούμε να χρησιμοποιήσουμε ένα νευρωνικό δίκτυο, προκειμένου να μοντελοποιήσουμε τη συνάρτηση  $F ()$  που υπάρχει στην παραπάνω εξίσωση. Η δομή του συγκεκριμένου νευρωνικού δικτύου αναπαριστάνεται από την εικόνα που ακολουθεί.



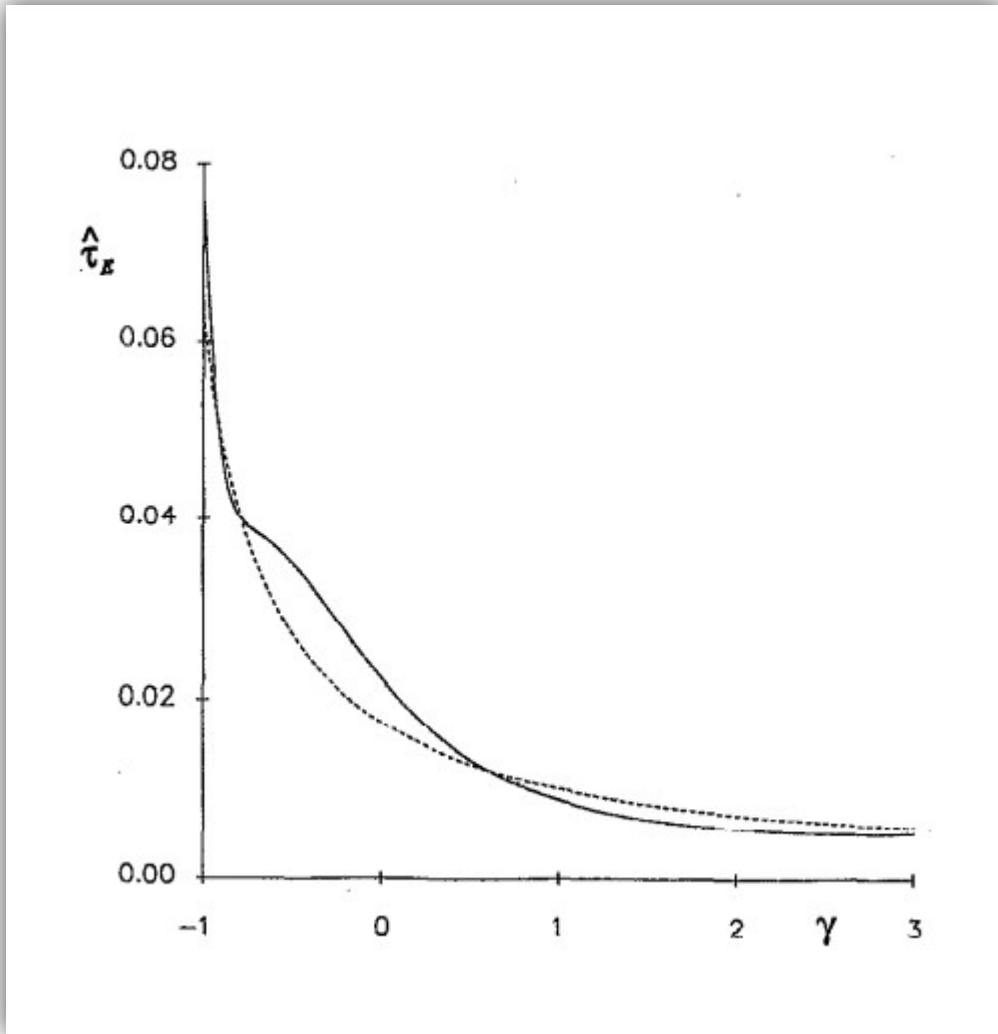
Εικόνα 11 - Δομή νευρωνικού δικτύου

Πρέπει να σημειωθεί ότι οι λογάριθμοι χρησιμοποιούνται τόσο ως μορφή προεπεξεργασίας των μεταβλητών εισόδου, όσο και για την προεπεξεργασία των δεδομένων στόχων, που είναι το  $\ln \hat{\tau}_E$ . Με τον τρόπο αυτό ελαττώνεται το δυναμικό εύρος των μεταβλητών και διασφαλίζεται ότι η σχετική ακρίβεια διατηρείται ακόμα και όταν ορισμένες από τις ποσότητες έχουν μικρές τιμές. Συμπεραίνουμε ότι εάν η χαρτογράφηση δικτύου είναι γραμμική, ανακτάται η τυπική έκφραση γραμμικής παλινδρόμησης. Το Νευρωνικό δίκτυο περιέχει την προσέγγιση γραμμικής παλινδρόμησης ως ειδική περίπτωση.

Οι δυσκολίες που πιθανόν να υπάρξουν, μπορούν να αποφευχθούν με τη βελτιστοποίηση της δομής του δικτύου. Αυτό επιτυγχάνεται συγκρίνοντας το τελικό δίκτυο με γραμμική παλινδρόμηση χρησιμοποιώντας ένα ξεχωριστό σύνολο δεδομένων δοκιμών.

Παρατηρείται μια μείωση του σφάλματος της τάξης του 25% όσον αφορά την προσέγγιση του νευρωνικού δικτύου, σε σύγκριση με τη γραμμική παλινδρόμηση.

Στην εικόνα που ακολουθεί παρατηρείται η σύγκριση της συμπεριφοράς για τη συνάρτηση  $F()$  που λαμβάνεται από το νευρωνικό δίκτυο, με το αντίστοιχο αποτέλεσμα από την προσέγγιση γραμμικής παλινδρόμησης.

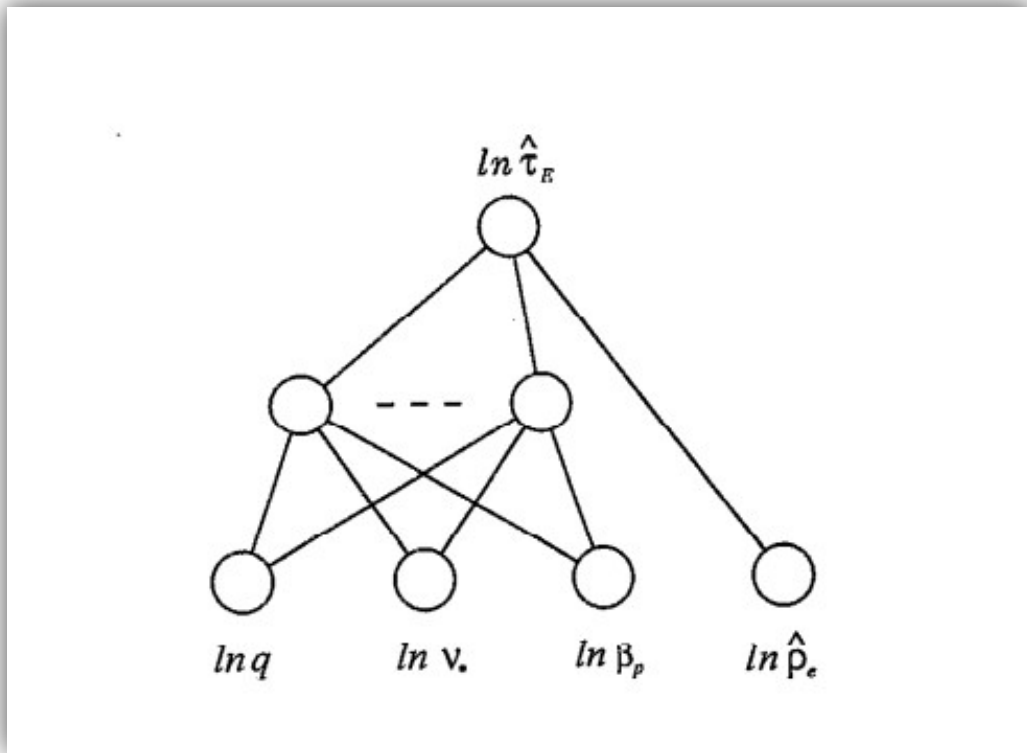


Εικόνα 12 - Σύγκριση αποτελεσμάτων για νευρωνικό δίκτυο με την παλινδρόμηση

Η συγκεκριμένη εφαρμογή προσφέρει μια εικόνα του τρόπου με τον οποίο η προηγούμενη γνώση μπορεί να ενσωματωθεί σε μια δομή νευρωνικού δικτύου. Η αναπαράσταση της  $\hat{\tau}_E$  μπορεί επίσης να έχει την ακόλουθη μορφή.

$$\hat{\tau}_E = (\hat{\rho}_e)^a G(\mathbf{q}, \mathbf{v}^*, \beta_p)$$

Στην περίπτωση όπου τα δεδομένα επεξεργαστούν και γίνεται πάλι χρήση λογαρίθμου, τότε η λειτουργική μορφή της παραπάνω εξίσωσης μπορεί να αναπαρασταθεί από τη δομή δικτύου που φαίνεται στην ακόλουθη εικόνα.



Εικόνα 13 - Τροποποιημένη έκδοση του δικτύου

Πρόκειται για μια τροποποιημένη μορφή του νευρωνικού δικτύου, στο οποίο το  $\hat{\tau}_E$  είναι μια συνάρτηση ισχύος της μεταβλητής εισόδου  $\hat{\rho}_e$ .

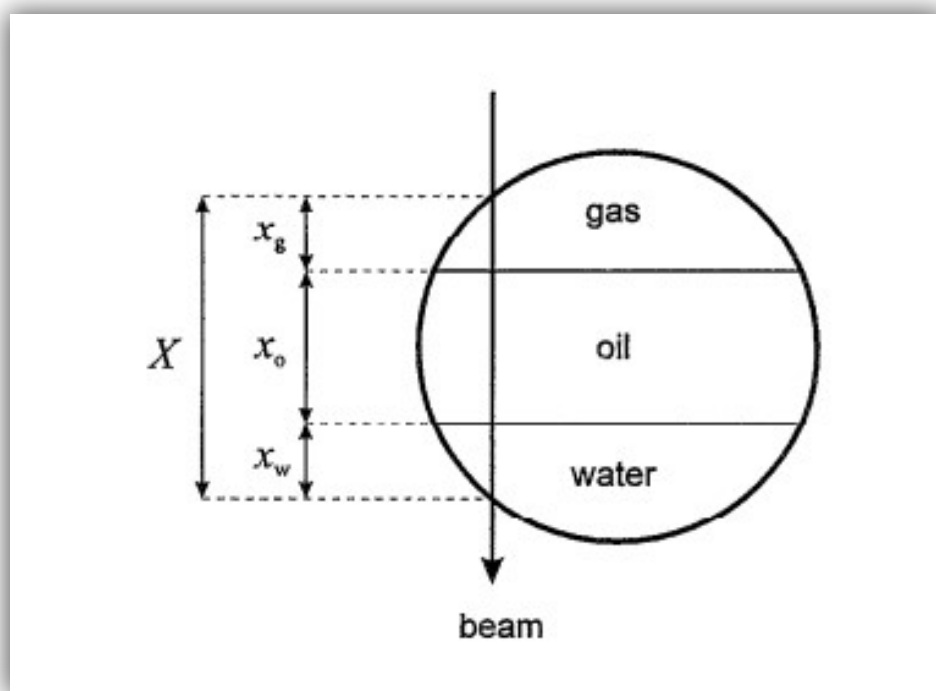
Ένα άλλο σημαντικό χαρακτηριστικό των νευρωνικών δικτύων όταν χρησιμοποιούνται για τη διεξαγωγή μη γραμμικής παρεμβολής είναι η ικανότητά τους να μάθουν πώς να συνδυάζουν δεδομένα από διάφορες αισθητήρες για να παράγουν σημαντικές εξόδους.

### 1.11.2 Classification

Το θέμα της συγκεκριμένης υποενότητας είναι μια εφαρμογή ταξινόμησης με νευρωνικά δίκτυα. Το συγκεκριμένο πρόβλημα που θα αναλυθεί, αναφέρεται στην παρακολούθηση της ροής πετρελαίου κατά μήκος των αγωγών που φέρουν μείγμα πετρελαίου, νερού και αερίου. Η επίλυση στοχεύει στην παροχή τεχνικής

για τη μέτρηση των ποσοστών ροής πετρελαίου. Το συγκεκριμένο πρόβλημα έχει πολύ μεγάλη σημασία για τη βιομηχανία πετρελαίου.

Η προσέγγιση βασίζεται στην τεχνική με ονομασία «dual-energy gamma densitometry». Γεγονός που περιλαμβάνει τη μέτρηση της εξασθένησης μίας ακτινικής δέσμης μονοενεργών gammas που διέρχεται μέσω του σωλήνα. Αυτό που μόλις αναφέρθηκε περιγράφεται καλύτερα από την εικόνα που ακολουθεί στη συνέχεια.



Εικόνα 14 - Σχηματική διατομή ενός αγωγού πετρελαίου

Για μια ακτίνα γάμμα που διέρχεται από μία μοναδική ομοιογενή ουσία το κλάσμα της έντασης της δέσμης  $I$  εξασθενημένο ανά μονάδα μήκους είναι σταθερό και έτσι η ένταση θα αποσυντεθεί εκθετικά με την απόσταση σύμφωνα με τον ακόλουθο τύπο.

$$I = I_0 e^{-\mu\rho x}$$

Για μια ακτίνα γάμμα που διέρχεται από ένα συνδυασμό πετρελαίου, νερού και αερίου, η ένταση της δέσμης διασπάται σύμφωνα με τη παρακάτω σχέση.

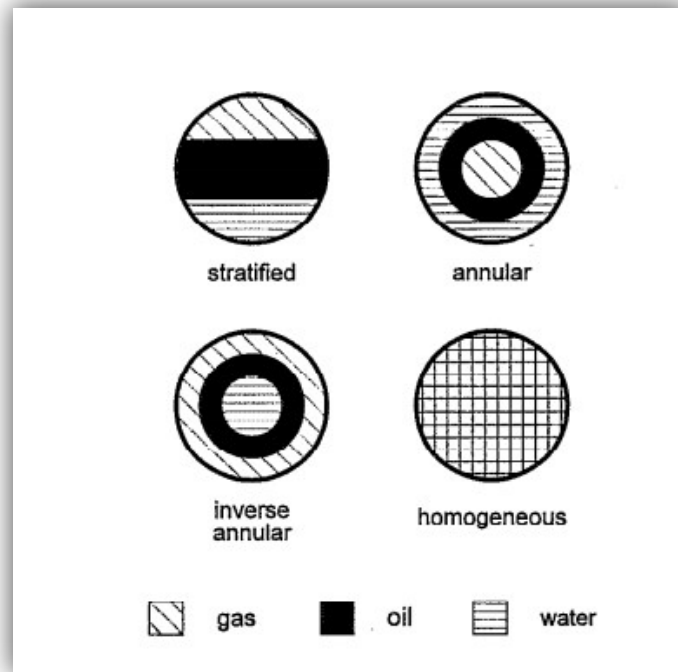
$$I = I_0 e^{-\mu_0 \rho_0 x_0} e^{-\mu_w \rho_w x_w} e^{-\mu_g \rho_g x_g}$$

Στην παραπάνω σχέση τα  $x_0$ ,  $x_w$ ,  $x_g$  αντιπροσωπεύουν τα μήκη της κάθε διαδρομής όπως φαίνεται και στην εικόνα που βρίσκεται παραπάνω. Η μέτρηση από μία γραμμή δέσμης δεν παρέχει επαρκείς πληροφορίες για να προσδιοριστούν και τα 3 μήκη διαδρομής. Για το λόγο αυτό, μια δεύτερη δέσμη γ της διαφορετικής ενέργειας περνάει κατά μήκος της ίδιας διαδρομής με την πρώτη δέσμη.

Τέλος, τα τρία μήκη διαδρομής είναι περιορισμένα για να προσθέσουν μέχρι το συνολικό μήκος διαδρομής μέσω του σωλήνα, όπως φαίνεται επίσης από την εικόνα που βρίσκεται παραπάνω.

$$x_0 + x_w + x_g = X$$

Στο σχήμα που ακολουθεί, αναπαριστώνται τέσσερις διαμορφώσεις μοντέλων των 3 φάσεων ροών που χρησιμοποιούνται για τη δημιουργία σύνθετων δεδομένων για τη μελέτη νευρωνικού δικτύου.



Εικόνα 15 - Τέσσερις διαμορφώσεις μοντέλων τριών φάσεων ροής

Το πρόβλημα μπορεί να αναλυθεί με τη χρήση των τεχνολογιών των νευρωνικών δικτύων. Πιο αναλυτικά, από το σύστημα 6 πυκνομέτρων διπλής ενέργειας μπορούμε να εξαγάγουμε τις αντίστοιχες 6 τιμές του μήκους διαδρομής στο λάδι και 6 τιμές του μήκους διαδρομής στο νερό. Αυτό δίνει 12 μετρήσεις από τις οποίες μπορούμε να προσπαθήσουμε να προσδιορίσουμε τη γεωμετρική διαμόρφωση.

Στη συνέχεια τα κλάσματα του πετρελαίου, του νερού και του αερίου για αυτή τη διαμόρφωση επιλέχθηκαν τυχαία με ομοιόμορφη κατανομή πιθανότητας, με την επιφύλαξη του περιορισμού που πρέπει να προσθέσουν στην ενότητα. Τα 12 ανεξάρτητα μήκη διαδρομής υπολογίζονται γεωμετρικά και αυτά σχηματίζουν τις εισόδους σε ένα νευρωνικό δίκτυο.

Η κυρίαρχη πηγή θορύβου στην εφαρμογή αυτή προκύπτει από τα στατιστικά στοιχεία φωτονίων και αυτά περιλαμβάνονται στα δεδομένα χρησιμοποιώντας τη σωστή κατανομή Poisson.

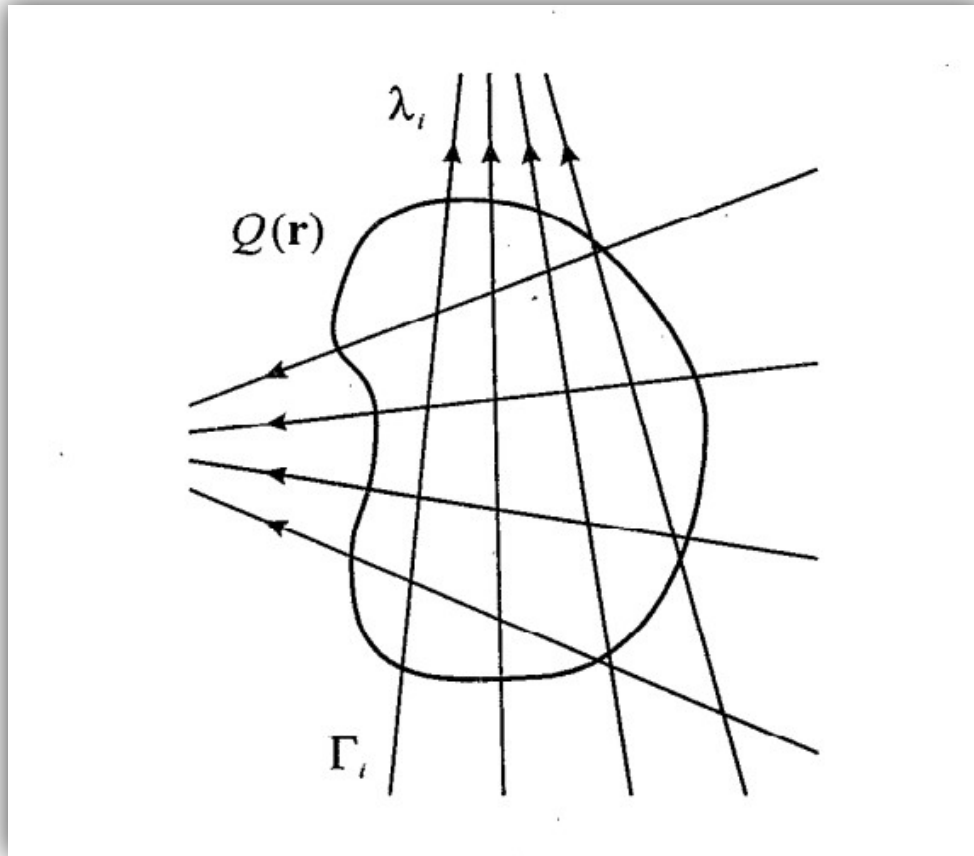
Τα δίκτυα εκπαιδεύτηκαν χρησιμοποιώντας ένα σύνολο δεδομένων 1.000 παραδειγμάτων και στη συνέχεια δοκιμάστηκαν χρησιμοποιώντας άλλα 1.000 ανεξάρτητα παραδείγματα. Η δομή δικτύου αποτελείται από ένα πολλαπλών επιπέδων perceptron, με ένα μόνο κρυφό και ένα επίπεδο εξόδου.

Συμπεραίνουμε ότι μια πιο άμεση προσέγγιση είναι να εκπαιδευτεί ένα δίκτυο για να χαρτογραφήσει τις πληροφορίες μήκους διαδρομής από τους πυκνομετρητές. Αυτό οδηγεί σε ένα δίκτυο με 12 εισόδους και 2 εξόδους που αντιστοιχούν στα κλάσματα όγκου του πετρελαίου και του νερού.

### 1.11.3 Inverse problems

Ένα επιπλέον είδος εφαρμογών που μπορούν να χρησιμοποιήσουν τα νευρωνικά δίκτυα είναι τα αντίστροφα προβλήματα. Με τη χρήση ενός παραδείγματος μπορούμε να κατανοήσουμε καλύτερα την συγκεκριμένη έννοια.

Στην εικόνα που ακολουθεί απεικονίζεται το πρόβλημα της γενικής τομογραφίας.



Εικόνα 16 - Το πρόβλημα της τομογραφίας

Ο στόχος είναι να προσδιοριστεί η κατανομή μιας ποσότητας  $Q(r)$  από έναν αριθμό μετρήσεων που γίνεται κατά μήκος διαφόρων οπτικών πεδίων μέσα από μια δεδομένη περιοχή. Ισχύει ο ακόλουθος τύπος.

$$\lambda_i = \int_{\Gamma_i} Q(r) dr$$

Η ποσότητα  $Q(r)$  θα μπορούσε να είναι μια απλή εκπομπή ακτινών  $X$  από ένα πλάσμα tokamak. Υπάρχουν και άλλα παραδείγματα, τα οποία περιλαμβάνουν τομογραφία απορρόφησης ακτίνων  $X$  σε ιατρικές εφαρμογές και υπερηχητική τομογραφία.



Στο συγκεκριμένο πρόβλημα που εξετάζετε θεωρούμε ότι είναι κατανομή  $Q(r)$  είναι γνωστή και θέλουμε να προβλέψουμε τις τιμές των ολοκληρωμένων γραμμών λι. Το πρόβλημα αυτό έχει μια μοναδική λύση που λαμβάνεται απλά από την αξιολόγηση των ολοκληρωμάτων της παραπάνω εξίσωσης. Ωστόσο, στην πράξη απαιτείται να λύσουμε το αντίστροφο πρόβλημα για τον προσδιορισμό της συνάρτησης  $Q(r)$  από έναν πεπερασμένο αριθμό μετρήσεων λι.

Επιπλέον, οι μετρήσεις ενδέχεται να αλλοιωθούν από θόρυβο. Επομένως μπορεί να αναζητήσουμε περισσότερες λύσεις που να ταιριάζουν ακριβώς στα δεδομένα. Πολλά από τα προβλήματα που προκύπτουν στην ανάλυση δεδομένων στα οποία μπορούν να εφαρμοστούν νευρωνικά δίκτυα είναι αντίστροφα προβλήματα.

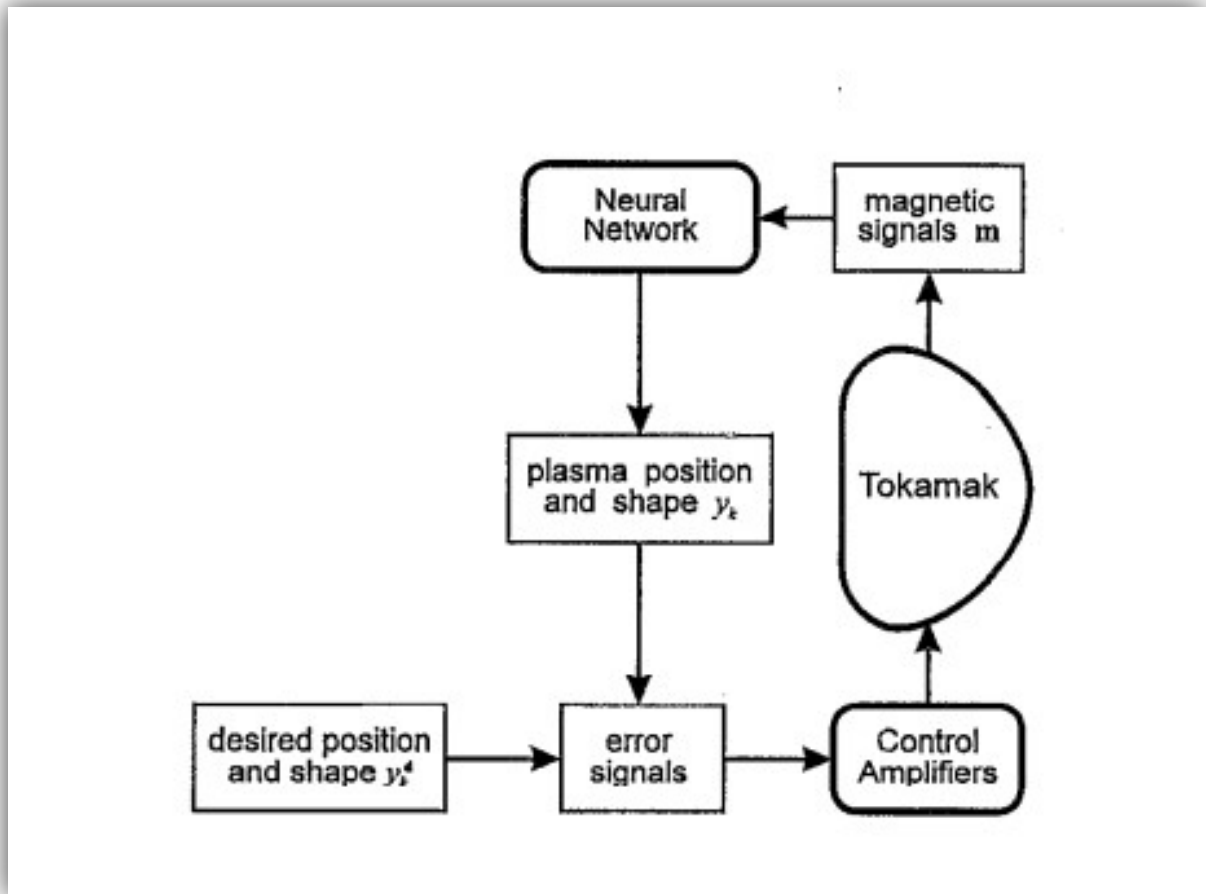
Μπορεί ένα πρόβλημα να έχει μια εύκολη λύση, ενώ το αντίστροφο πρόβλημα μπορεί να είναι δύσκολο και να απαιτεί πολλές προσεγγίσεις και υπολογιστικές επαναλήψεις μέχρις ότου βρεθεί η λύση. Η προσέγγιση του νευρωνικού δικτύου προσφέρει το σημαντικό πλεονέκτημα της πολύ υψηλής ταχύτητας.

#### 1.11.4 Control applications

Αρχικά, πρέπει να σημειωθεί ότι τα νευρωνικά δίκτυα παρέχουν λύσεις σε ένα πολύ μεγάλο βαθμό πολύπλοκων προβλημάτων όσον αφορά τον μη γραμμικό έλεγχο. Στη συγκεκριμένη υποενότητα αναφερόμαστε στον κλάδο των νευρωνικών δικτύων για την ανάλυση των δεδομένων.

Τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (Feedforward Networks), τα οποία έχουν αναφερθεί και έχουν περιγραφεί αναλυτικότερα σε προηγούμενη υποενότητα, μπορούν να χρησιμοποιηθούν για την εκτέλεση μη γραμμικής χαρτογράφησης στο πλαίσιο ενός κλασικού γραμμικού κυκλώματος ελέγχου ανατροφοδότησης.

Στην εικόνα που βρίσκεται στη συνέχεια, αναπαριστάνεται η αξιοποίηση της συγκεκριμένης τεχνικής με μεγάλη επιτυχία για τον έλεγχο της ανατροφοδότησης του tokamak.



Εικόνα 17 - Επιτυχής έλεγχος ανατροφοδότησης του tokamak

Ο κλάδος των νευρικών δικτύων χρησιμοποιήθηκε για τον έλεγχο ανατροφοδότησης σε πραγματικό χρόνο. Σκοπό είχε να ελέγχει τη θέση και το σχήμα του πλάσματος στο πείραμα TOAMAK της COMPASS με χρήση του συστήματος ελέγχου.

Ως εισόδους στο συγκεκριμένο δίκτυο υπάρχει ένα σύνολο μαγνητικών σημάτων  $m$ , τα οποία περιβάλλουν το tokamak. Τα σήματα αυτά λαμβάνονται από ένα σύνολο πηνίων, που βρίσκεται γύρω από το tokamak. Εν συνεχεία, το νευρωνικό δίκτυο χαρτογραφεί τα μαγνητικά σήματα, παρουσιάζοντας ένα σύνολο γεωμετρικών παραμέτρων. Οι παράμετροι αυτοί αντιστοιχούν στη θέση καθώς και στο σχήμα που οριοθετεί το πλάσμα. Οι τιμές που έχουν αυτές οι παράμετροι από το νευρωνικό δίκτυο στη συνέχεια θα συγκριθούν με τις επιθυμητές τιμές.

Τα σήματα σφάλματος που προκύπτουν, έπειτα στέλνονται στους τυπικούς ενισχυτές γραμμικού ελέγχου. Με τη σειρά τους οι ενισχυτές αυτοί ρυθμίζουν τη θέση και το σχήμα του πλάσματος. Η ρύθμιση πραγματοποιείται με τη μεταβολή του αριθμού των πηνίων ελέγχου.

Το συγκεκριμένο νευρωνικό δίκτυο που μελετάται στην ενότητα αυτή εκπαιδεύεται χωρίς σύνδεση στο λογισμικό. Με τη χρήση ενός μεγάλου συνόλου δεδομένων από παραδειγματικές διαμορφώσεις πλάσματος. Ένα τέτοιου είδους σύστημα πέτυχε πρόσφατα τον πρώτο έλεγχο σε πραγματικό χρόνο ενός πλάσματος tokamak από ένα νευρωνικό δίκτυο.

Η συγκεκριμένη εφαρμογή παρέχει μια επιπλέον ιδιότητα. Πρόκειται για τη συνέπεια της γραμμικότητας των εξισώσεων του Maxwell. Συγκεκριμένα, εάν όλα τα ρεύματα στο σύστημα tokamak κλιμακωθούν με έναν σταθερό παράγοντα, οι τιμές του μαγνητικού πεδίου θα κλιμακωθούν με τον ίδιο σταθερό παράγοντα καθώς και η θέση και το σχήμα του πλάσματος θα παραμείνουν αμετάβλητες. Πιο αναλυτικά, συμπεράνουμε ότι η χαρτογράφηση από τα σήματα στις παραμέτρους τόσο της θέσης όσο και του σχήματος του πλάσματος θα πρέπει να έχει την συγκεκριμένη ιδιότητα. Δηλαδή, εάν όλες οι εισοδοί του δικτύου κλιμακωθούν με τον ίδιο συντελεστή θα πρέπει και οι έξοδοι του να παραμείνουν αμετάβλητες.

Στο σημείο αυτό, είναι σκόπιμο να αναφέρουμε μια επιπλέον πρόσφατη εφαρμογή των Νευρωνικών δικτύων που πραγματοποιήθηκε σε πραγματικό χρόνο. Η συγκεκριμένη εφαρμογή ήταν ο έλεγχος 6 τμημάτων καθρέπτη σε ένα αστρονομικό οπτικό τηλεσκόπιο. Ο έλεγχος αυτός πραγματοποιήθηκε με σκοπό να υλοποιείται σε πραγματικό χρόνο ακύρωσης των στρεβλώσεων που πιθανόν να προκληθούν από διάφορες ατμοσφαιρικές αναταράξεις.

Η συγκεκριμένη εφαρμογή που αναφέρθηκε προηγουμένως ονομάζεται προσαρμοστική οπτική - adaptive optics. Πρόκειται για μια εφαρμογή, η οποία προκαλεί αλλαγή στο σχήμα του καθρέπτη κάθε 10 ms. Οι συμβατικές προσεγγίσεις περιλαμβάνουν επαναληπτικό αλγόριθμο. Το νευρωνικό δίκτυο παρέχει μια γρήγορη εναλλακτική λύση, η οποία επιτυγχάνει υψηλή ακρίβεια.

Εκτός από τα όσα αναφέρθηκαν προηγουμένως, πρέπει να σημειωθεί ότι τα νευρωνικά δίκτυα μπορούν επιπλέον να χρησιμοποιηθούν ως μη γραμμικά προσαρμοστικά εξαρτήματα μέσα σε ένα βρόχο ελέγχου. Το νευρωνικό δίκτυο στην συγκεκριμένη περίπτωση συνεχίζει να εκπαιδεύεται ενώ ενεργεί ως ελεγκτής. Επίσης, το δίκτυο αυτό μπορεί να εκπαιδεύεται ώστε να ελέγχει πολύπλοκα μη γραμμικά συστήματα με δοκιμές και λάθη.

# ΚΕΦΑΛΑΙΟ 2 – ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ

---

## 2.1 Εισαγωγή – Βασικές έννοιες

Ένα ιδιαίτερα σημαντικό χαρακτηριστικό στη συμπεριφορά των ανθρώπων είναι η ικανότητά τους στη Μάθηση (Learning). Στο γεγονός αυτό βασίστηκαν οι επιστήμονες του κλάδου των Τεχνητών Νευρωνικών Δικτύων. Πιο συγκεκριμένα, οι επιστήμονες κατασκευάζουν υπολογιστικά συστήματα, με σκοπό να μπορούν να μαθαίνουν να εκπαιδεύονται και να επιτύχουν τους προβλεπόμενους στόχους μέσω της Μηχανικής Μάθησης (Machine Learning). (Haykin, 2009)

Ένας ορισμός που μπορεί να δοθεί για τη Μηχανική Μάθηση είναι ότι αποτελεί ένα φαινόμενο, στο οποίο τα συστήματα αυξάνουν συνεχώς την απόδοση τους εκτελώντας μια διεργασία ενώ δεν χρειάζεται να προγραμματίζονται ξανά από την αρχή. (Αργυράκης, 2001)

Αντικείμενο του κλάδου της Μηχανικής Μάθησης είναι η δημιουργία κατάλληλων προγραμμάτων, τα οποία θα έχουν την ικανότητα να βελτιώνουν την απόδοσή τους αυτόνομα με βάση τις γνώσεις και την εμπειρία που διαθέτουν. Πρόκειται για έναν συνεχώς εξελισσόμενο κλάδο της επιστήμης. Η μηχανική μάθηση ασχολείται κατά κύριο λόγο με αλγορίθμους βελτιστοποίησης και αλληλεπιδρά με το γνωστικό πεδίο διάφορων επιστημονικών κλάδων. Με τη χρήση των αλγορίθμων μάθησης επιτυγχάνεται η κατάλληλη εξαγωγή γενικευμένων κανόνων μάθησης, προκειμένου να επιλύονται ορισμένα προβλήματα με βάση ένα πεπερασμένο σύνολο πληροφοριών εκπαίδευσης.

Οι Mendel και McClaren αναφέρονται στη μάθηση όσον αφορά τα νευρωνικά δίκτυα ως τη διαδικασία κατά την οποία αναδιαμορφώνονται οι παράμετροι του δικτύου. Η συγκεκριμένη διαδικασία είναι μια συνεχής διέγερση του δικτύου από το περιβάλλοντος του. Ο τρόπος διαμόρφωσης των παραμέτρων ενός δικτύου καθορίζει κάθε φορά το είδος της εκπαίδευσής του. (Πλέρου, 2012)

Όλα τα διαφορετικά είδη μάθησης που έχουν δημιουργηθεί από τους επιστήμονες του τομέα των νευρωνικών δικτύων από την αρχή της ανάπτυξής τους, έχουν ελεγχθεί και εφαρμοστεί με μεγάλη επιτυχία σε πολλές λειτουργίες διαφορετικών τομέων. Ως παραδείγματα τέτοιων λειτουργιών μπορούν να αναφερθούν οι προβλέψεις σε οικονομικά θέματα, η ταξινόμηση κειμένων κτλ.

## 2.2 Εκπαίδευση Νευρωνικών Δικτύων

Όπως έχει ήδη αναφερθεί, ο κύριος στόχος που προέτρεψε τους επιστήμονες να κατασκευάσουν τα τεχνητά νευρωνικά δίκτυα ήταν η δημιουργία συστημάτων, τα οποία θα μπορούσαν να επιλύσουν διάφορες διεργασίες καθώς και να εκτελούν από μόνα τους μια λειτουργία. Προκειμένου να συμβεί κάτι ανάλογο θα πρέπει το δίκτυο να εκπαιδευτεί σωστά.

Στα τεχνητά Νευρωνικά Δίκτυα παρατηρούνται δύο βασικές ιδιότητες. Αυτές οι ιδιότητες είναι η μάθηση (learning) και η ανάκληση (recall). Με τον όρο Ανάκληση, αναφερόμαστε στη διαδικασία κατά την οποία πραγματοποιείται υπολογισμός της εξόδου με δεδομένη την είσοδο και τις τιμές από τα βάρη. Με τον όρο Μάθηση Νευρωνικών Δικτύων αναφερόμαστε στη διαδικασία κατά την οποία πραγματοποιούνται ορισμένες τροποποιήσεις στις τιμές των βαρών, με σκοπό να παραχθεί μια συγκεκριμένη, προκαθορισμένη έξοδος με δεδομένη την είσοδο του δικτύου. (Haykin, 2009)

Για να πραγματοποιηθεί σωστά η διαδικασία της μάθησης ενός δικτύου είναι απαραίτητη η χρησιμοποίηση ορισμένων παραδειγμάτων καθώς και ενός αλγορίθμου εκπαίδευσης. Η διαδικασία της μάθησης γίνεται αποκλειστικά για να βελτιωθεί η απόδοση ενός τεχνητού νευρωνικού δικτύου. Ο αλγόριθμος μάθησης εκτελεί επαναληπτικές λειτουργίες, τροποποιώντας κατάλληλα τις τιμές των βαρών και του κατωφλιού ενός δικτύου. Με τη χρήση του συγκεκριμένου αλγορίθμου ελατώνονται σημαντικά τα σφάλματα που υπάρχουν ανάμεσα στη πραγματική έξοδο και την επιθυμητή έξοδο κάθε δικτύου. Η διαδικασία της μάθησης αναφέρεται και ως διαδικασία εκπαίδευσης. Στα τεχνητά νευρωνικά δίκτυα διακρίνονται πολλά είδη μάθησης των δικτύων. Το σημείο που διαφέρουν όμως είναι ο τρόπος με τον οποίο τροποποιούνται οι παράμετροι του δικτύου που εκπαιδεύεται.

Η μάθηση ενός δικτύου διακρίνεται από τα βήματα που ακολουθούν:

- **Διέγερση** του δικτύου από το εξωτερικό περιβάλλον.
- Ως αποτέλεσμα της διέγερσης του δικτύου από το περιβάλλον είναι οι **αλλαγές** που υφίσταται στις παραμέτρους του.
- Το τεχνητό νευρωνικό δίκτυο που υφίσταται τις αλλαγές στο εσωτερικό του, **αντιδράει** σύμφωνα με έναν καινούργιο τρόπο στο περιβάλλον του. (D. Anderson, 1992)

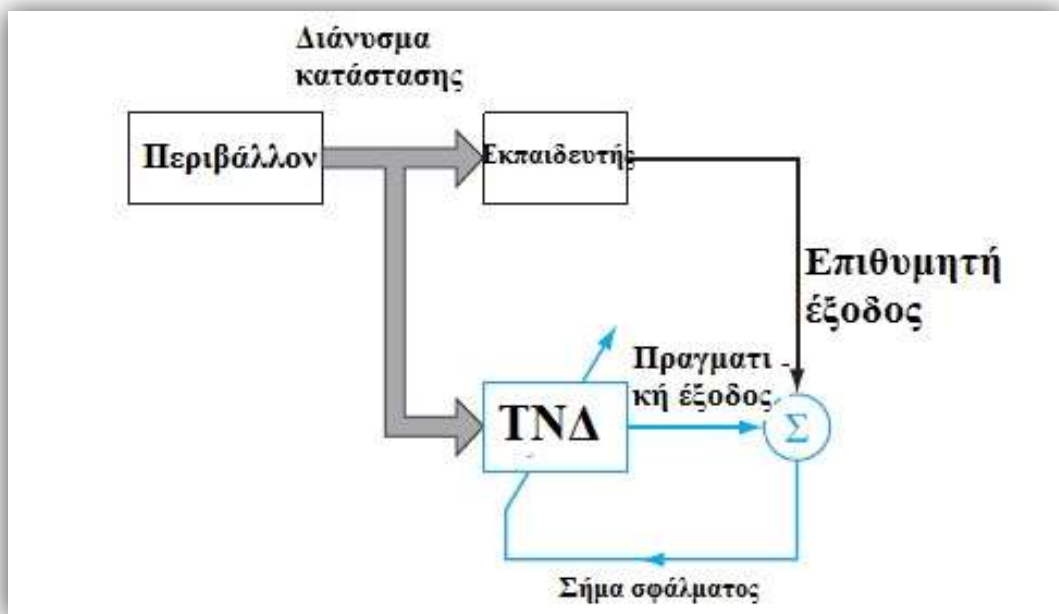
## 2.3 Είδη Μηχανικής Μάθησης

Όπως ήδη αναφέρθηκε, υπάρχουν διαφορετικές μορφές μάθησης των νευρωνικών δικτύων. Στη συνέχεια αυτού του κεφαλαίου αναφέρονται και αναλύονται τα τρία είδη μηχανικής μάθησης στα τεχνητά νευρωνικά δίκτυα, της επιβλεπόμενης, της ενισχυτικής και μη επιβλεπόμενης μάθησης.

### 2.3.1 Επιβλεπόμενη Μάθηση (Supervised Learning)

Στο συγκεκριμένο είδος μάθησης με επίβλεψη το νευρωνικό δίκτυο λαμβάνει ζεύγη από εισόδους και επιθυμητές εξόδους. Με δεδομένες τις τιμές των βαρών το κάθε δίκτυο παράγει μια έξοδο. Στην έξοδο αυτή παρατηρούνται διαφορές ως προς την επιθυμητή έξοδο. Πιο συγκεκριμένα, παρατηρείται αρχικά ένα σφάλμα (error) της εξόδου του δικτύου με την επιθυμητή έξοδο. Σύμφωνα με το σφάλμα αυτό καθώς και ενός αλγορίθμου εκπαίδευσης πραγματοποιούνται τροποποιήσεις στις τιμές των βαρών.

Μπορούμε να αναφέρουμε τη μάθηση με επίβλεψη και ως *μάθηση με εκπαιδευτή*. Αυτός ο τρόπος μάθησης αποτελεί τον πιο συχνά χρησιμοποιούμενο τρόπο για την εκπαίδευση των Τεχνητών νευρωνικών δικτύων. Στην Εικόνα 11 που βρίσκεται παρακάτω παρουσιάζεται το διάγραμμα της διαδικασίας για την επιβλεπόμενη μάθηση. (Haykin, 2009)



Εικόνα 18 - Διάγραμμα επιβλεπόμενης μάθησης (Supervised Learning)

Στο διάγραμμα της Εικόνας 11 παρατηρείται ότι η επιβλεπόμενη μάθηση είναι ένα σύστημα, το οποίο παρουσιάζει ανάδραση κλειστού βρόχου. Το περιβάλλον του δικτύου επίσης βρίσκεται μέσα στο βρόχο. Το άθροισμα των τετραγώνων όλων των σφαλμάτων ή το μέσο τετραγωνικό σφάλμα συναρτηθεί των τιμών όλων των βαρών του δικτύου αποτελεί το μέτρο απόδοσης του συγκεκριμένου δικτύου. (Haykin, 2009). Δεν είναι λίγες οι φορές κατά τις οποίες οι τροποποιήσεις των τιμών στα βάρη του δικτύου να είναι πάρα πολλές. Αυτό έχει ως αποτέλεσμα να αυξάνονται αρκετά οι υπολογιστικοί χρόνοι.

Το Τεχνητό Νευρωνικό Δίκτυο δεν κατέχει γνώσεις και πληροφορίες για το περιβάλλον του, εν αντιθέσει με τον εκπαιδευτή του. Οι γνώσεις αυτές προκύπτουν μέσα από ένα πλήθος παραδειγμάτων εισόδου – εξόδου. Ένα δiάνυσμα εκπαίδευσης (παραδειγμα εισόδου ) εισέρχεται από το περιβάλλον στο Νευρωνικό δίκτυο. Ο εκπαιδευτής με τη σειρά του, με βάση το συγκεκριμένο δiάνυσμα εισόδου εκπαίδευσης παράγει μια επιθυμητή απόκριση για το δίκτυο. Με τον όρο επιθυμητή απόκριση αναφερόμαστε στη καλύτερη δυνατή ενέργεια που απαιτείται από το Νευρωνικό Δίκτυο. Όπως έχει ήδη αναφερθεί το σφάλμα σήματος είναι η διαφορά που υπάρχει ανάμεσα στην πραγματική και επιθυμητή απόκριση. Οι τιμές των βαρών του δικτύου μεταβάλλονται διαρκώς με επαναληπτική μέθοδο. Η μέθοδος αυτή στοχεύει στο να βρεθεί το Νευρωνικό Δίκτυο σε μια συγκεκριμένη κατάσταση. Σε αυτήν την κατάσταση το δίκτυο θα προσομοιώνει όσο το δυνατόν καλύτερα τη συμπεριφορά του εκπαιδευτή. (Russell Reed, 1999)

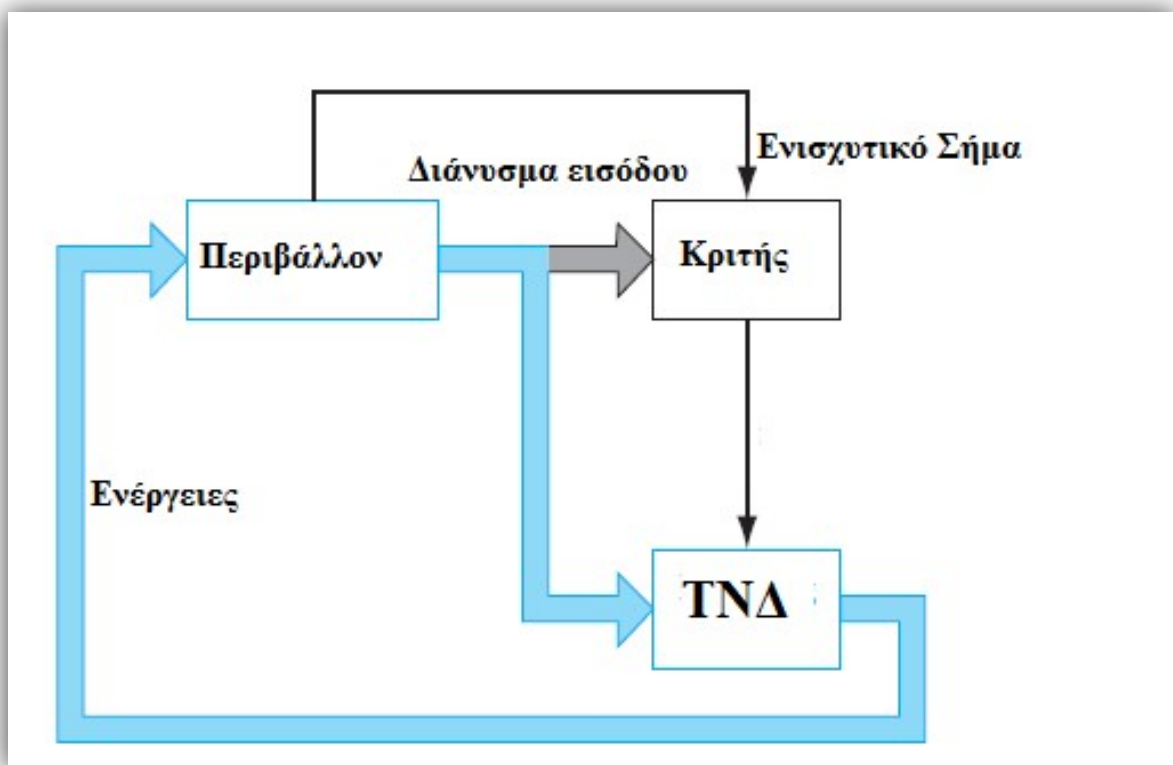
Τα συναπτικά βάρη των Νευρωνικών δικτύων αποτελούν τη μνήμη του δικτύου, στα οποία αποθηκεύονται οι γνώσεις και οι πληροφορίες. Με τη διαδικασία που περιγράφηκε προηγουμένως, οι γνώσεις από το περιβάλλον μέσω της εκπαίδευσης φτάνουν στο Νευρωνικό Δίκτυο. Οι γνώσεις αυτές αποθηκεύονται στα βάρη του δικτύου ως σταθερές.

### **2.3.2 Ενισχυτική Μάθηση (Reinforcement Learning)**

Η δεύτερη μέθοδος μάθησης που αναλύεται στη συνέχεια είναι η ενισχυτική μάθηση (Reinforcement learning). Στη συγκεκριμένη μέθοδο η έξοδος κάθε δικτύου μπορεί να είναι είτε καλή είτε κακή. Ο χαρακτηρισμός αυτός καθορίζεται σύμφωνα με μια αριθμητική κλίμακα. Τα βάρη αναδιαμορφώνονται βάση του συγκεκριμένου χαρακτηρισμού. (W. Thomas Miller, 1990). Τα Τεχνητά Νευρωνικά δίκτυα που εκπαιδεύονται με ενισχυτική μάθηση τροφοδοτούνται με δείγματα εισόδων. Στα πρότυπα εισόδων δεν υφίσταται οι επιθυμητές αποκρίσεις. Σε αυτόν τον τύπο μάθησης γίνεται χρήση ενός συνολικού μέτρου

επάρκειας του δείκτη συμπεριφοράς. Αυτό το μέτρο αναφέρεται ως ενισχυτικό σήμα (reinforcement signal) και έχει την ικανότητα να καθοδηγεί το νευρωνικό δίκτυο κάθε φορά στην επιθυμητή συμπεριφορά. Αυτό επιτυγχάνεται, τροφοδοτώντας εκ νέου το δίκτυο και επιβραβεύοντας (reward) τις σωστές, κατάλληλες συμπεριφορές και τιμωρώντας τις λάθος συμπεριφορές (penalize). (Sutton, 1998)

Στην Εικόνα 12 που βρίσκεται παρακάτω παρουσιάζεται το διάγραμμα της διαδικασίας για την ενισχυτική μάθηση. (Haykin, 2009) Ο συγκεκριμένος τύπος μάθησης βασίζεται στο νόμο *Thorndike*. Σύμφωνα με το νόμο αυτό μια ενέργεια που έχει μια ικανοποιητική συμπεριφορά μέσα στη διαδικασία μάθησης, έχει ως αποτέλεσμα να ενισχύεται η τάση του δικτύου να παράγει τη συγκεκριμένη ενέργεια. Εν αντιθέσει με μια ενέργεια η οποία δεν διαθέτει μια ικανοποιητική συμπεριφορά, έχει ως αποτέλεσμα να εξασθενεί συνεχώς η τάση του δικτύου που εκπαιδεύεται να την παράγει.



Εικόνα 19 - Ενισχυτική Μάθηση (Reinforcement Learning)



Διακρίνονται δύο διαφορετικοί τύποι της ενισχυτικής μάθησης. Ο πρώτος τύπος είναι η συσχετική ενισχυτική μάθηση. Σύμφωνα με το τύπο αυτό, εκτός από το ενισχυτικό σήμα που τροφοδοτείται στο Νευρωνικό Δίκτυο, τροφοδοτούνται επίσης διάφοροι τύποι πληροφοριών. Από τις πληροφορίες αυτές το δίκτυο πρέπει να απεικονίσει σύμφωνα με μια συσχέτιση αιτία (stimulus) – αποτέλεσμα (action). Ο δεύτερος τύπος είναι η μη συσχετιστική ενισχυτική μάθηση. Στον τύπο αυτό, η μόνη είσοδος από το περιβάλλον στο νευρωνικό δίκτυο είναι το ενισχυτικό σήμα. Αποτελεί τη μοναδική πληροφορία που παρέχεται στο δίκτυο. Εδώ δεν υφίσταται οι συσχετίσεις των ενεργειών με διαφορετικές εισόδους. Η μέθοδος αυτή στοχεύει στο να βρεθεί η μοναδική βέλτιστη ενέργεια.

### 2.3.3 Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)

Μια άλλη μέθοδος μάθησης των Τεχνητών Νευρωνικών Δικτύων είναι η μη επιβλεπόμενη μάθηση. Στη συγκεκριμένη μέθοδο, η απόκριση του Νευρωνικού Δικτύου καθορίζεται από την ιδιότητα του ίδιου να οργανώνεται σύμφωνα με τα διανύσματα εισόδου του. Ο λόγος που συμβαίνει αυτό είναι διότι οι αντίστοιχοι έξοδοι δεν υπάρχουν. Το κάθε δίκτυο οργανώνεται εσωτερικά έτσι ώστε ένας συγκεκριμένος νευρώνας να αντιδρά κάθε φορά σε μια συγκεκριμένη είσοδο. Δεν υφίσταται στη μη επιβλεπόμενη μάθηση ένας εξωτερικός εκπαιδευτής, ο οποίος να καθορίζει συγκεκριμένα την επιθυμητή απόκριση καθώς και να επιβλέπει τη μάθηση του δικτύου. Στη μέθοδο αυτή απλώς παρέχεται η πληροφορία και η γνώση στο Νευρωνικό δίκτυο. (Geoffrey E. Hinton, 1999)

Στην Εικόνα 13 που βρίσκεται παρακάτω παρουσιάζεται το διάγραμμα της διαδικασίας για τη μη επιβλεπόμενη μάθηση. (Haykin, 2009) Στη συγκεκριμένη μέθοδο, μόνο τα διανύσματα εισόδου αποτελούν τα δείγματα μάθησης. Δεν υφίσταται δείγματα για την επιθυμητή έξοδο. Στο Νευρωνικό δίκτυο υπάρχει ένας έλεγχος εσωτερικός. Σύμφωνα με τον έλεγχο αυτό, το δίκτυο αναζητά κανονικότητα στην είσοδο και εκπαιδεύεται, με σκοπό να κατασκευάσει εξόδους με τις ιδιότητες που έχουν και οι αντίστοιχες εισοδοί του.



**Εικόνα 20 - Διάγραμμα μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**

Χρησιμοποιώντας έναν μηχανισμό, ο οποίος ονομάζεται μηχανισμός ανάδρασης (feedback) το Νευρωνικό Δίκτυο ελέγχει διαρκώς την δράση του, εντοπίζει και διορθώνει τυχόν σφάλματα που υπάρχουν στα δεδομένα και τις πληροφορίες. Η εκπαίδευση τερματίζει όταν στο δίκτυο σταματήσουν οι τροποποιήσεις στις τιμές των βαρών του. (Haykin, 2009)

## **2.4 Λειτουργίες Μάθησης**

Στη συγκεκριμένη ενότητα αναφέρονται μερικές από τις βασικότερες εργασίες μηχανικής μάθησης στα τεχνητά νευρωνικά δίκτυα. Οι εργασίες μάθησης επηρεάζουν σε πολύ μεγάλο βαθμό την επιλογή ενός κανόνα μάθησης.

### **2.4.1 Συσχέτιση προτύπων**

Ένα ιδιαίτερα σημαντικό χαρακτηριστικό που παρατηρείται στη μνήμη των ζωντανών οργανισμών είναι η ικανότητα συσχέτισης. Έτσι και στα νευρωνικά δίκτυα η διαδικασία συσχέτισης αποτελεί χρήσιμη λειτουργία. Η συσχετική μνήμη που υπάρχει στα δίκτυα έχει τη μορφή μιας κατανεμημένης μνήμης. Μια τέτοια μορφή μνήμης, η οποία έχει τη δυνατότητα να μαθαίνει και να εκπαιδεύεται με τη βοήθεια συσχετίσεων είναι και ο ανθρώπινος εγκέφαλος. (Πλέρου, 2012)

Στα νευρωνικά δίκτυα διακρίνονται δύο μορφές συσχετίσεων. Η αυτό-συσχέτιση αποτελεί την πρώτη μορφή, σύμφωνα με την οποία κάθε δίκτυο πρέπει να δεσμεύει και να αποθηκεύει ένα πλήθος από πρότυπα που χρησιμοποιούνται με μεγάλη συχνότητα από το συγκεκριμένο δίκτυο. Σκοπός είναι όταν το δίκτυο βρει μια πληροφορία παρόμοια με κάποιο από τα αποθηκευμένα πρότυπα, να έχει την ικανότητα να το αναγνωρίζει ως πρότυπο και να το ανακαλέσει από τη μνήμη του. Η δεύτερη μορφή συσχετίσεων είναι η ετερο-συσχέτιση. Η ετερο-συσχέτιση διαφέρει σημαντικά από την αυτό-συσχέτιση. Πιο συγκεκριμένα, στην ετερο-συσχέτιση συνδυάζονται ένα τυχαίο πλήθος προτύπων εισόδων με ένα επίσης τυχαίο πλήθος προτύπων εξόδων. Στην αυτό-συσχέτιση η μορφή μάθησης που χρησιμοποιείται είναι η μη επιβλεπόμενη μάθηση. Εν αντιθέσει, με την ετερο-συσχέτιση, στην οποία χρησιμοποιείται ως εκπαίδευση η επιβλεπόμενη μάθηση. (Αργυράκης, 2001)

#### **2.4.2 Αναγνώριση προτύπων (pattern recognition)**

Όλοι οι ζωντανοί οργανισμοί, όπως είναι οι άνθρωποι και τα υπόλοιπα όντα, διαθέτουν την δυνατότητα να αναγνωρίζουν πραγματικές πληροφορίες, με τη βοήθεια διάφορων αισθήσεων και την ικανότητα της αντίληψης που έχουν. Με τον όρο Αναγνώριση Προτύπων (Pattern recognition) αναφερόμαστε σε έναν επιστημονικό κλάδο, ο οποίος στοχεύει στην απόδοση συγκεκριμένης τιμής στα διανύσματα εισόδου. Πρόκειται ουσιαστικά για μια διαδικασία κατά την οποία ένα σήμα που εισέρχεται στο δίκτυο ταυτίζεται με ένα πρότυπο από ένα προκαθορισμένο πλήθος κλάσεων. (Bishop, Neural Networks for Pattern Recognition, 1995)

Από τη δεκαετία κιόλας του 1960, που αναπτύχθηκε και η επιστήμη του υπολογιστή, οι επιστήμονες ασχολήθηκαν με τον τομέα της Αναγνώρισης των προτύπων. Αρχικά, όλες οι έρευνες αυτού του πεδίου στην εξέλιξη των θεωρητικών μεθόδων. Στην πορεία και συγκεκριμένα το 1970 οι προσπάθειες για το πεδίο αυτό συνεχώς εξελίσσονται. Η Παγκόσμια Ένωση Αναγνώρισης των Προτύπων (IARP) δημιουργήθηκε μόλις το 1976. (Bishop, Pattern Recognition and Machine Learning, 2006)

Κάθε μηχανή πρέπει να εκπαιδεύεται σωστά, με σκοπό να έχει τη δυνατότητα αναγνώρισης των προτύπων και κατ' επέκταση την ταξινόμησή τους αυτόματα σε διάφορες κατηγορίες. Μια τέτοια μηχανή είναι και ο ηλεκτρονικός υπολογιστής. Με τη χρήση κατάλληλων αλγορίθμων ταξινόμησης πραγματοποιείται η κατηγοριοποίηση των προτύπων σε κλάσεις.

Στα περισσότερα πεδία της επιστήμης χρησιμοποιούνται κατά κόρον οι εφαρμογές για την αναγνώριση των προτύπων. Ως παράδειγμα μπορεί να

αναφερθεί η επιστήμη της Ιατρικής και οι εφαρμογές της αναγνώρισης προτύπων που είναι η διάγνωση μέσω ηλεκτρονικού υπολογιστή και διάφορες εφαρμογές της βιο πληροφορικής. Ένα δεύτερο παράδειγμα που μπορεί να αναφερθεί είναι η επιστήμη των υπολογιστών. Ως εφαρμογές της αναγνώρισης προτύπων στο συγκεκριμένο επιστημονικό κλάδο εκτός από τα νευρωνικά δίκτυα και τη μηχανική μάθηση μπορούν να αναφερθούν επίσης η αναγνώριση χαρακτήρων και φωνής και τα συστήματα υποστήριξης αποφάσεων. Στον τομέα της βιομηχανίας επίσης χρησιμοποιούνται εφαρμογές αναγνώρισης προτύπων με σκοπό να επιτύχουν την αυτόματη επεξεργασία των δεδομένων. (Ripley, 1996)

Με την ολοκλήρωση μιας διαδικασίας μάθησης, τα ανθρώπινα όντα είναι σε θέση να εκτελέσουν την αναγνώριση των προτύπων. Το ίδιο ακριβώς ισχύει και για τα τεχνητά νευρωνικά δίκτυα. Υπάρχουν δύο διαφορετικές μορφές μηχανών για την αναγνώριση προτύπων στα νευρωνικά δίκτυα. Στη συνέχεια αναφέρονται και αναλύονται οι δύο αυτές μορφές.

- Στην πρώτη μορφή, το πρόγραμμα ή η μηχανή διασπάται σε δύο διαφορετικά τμήματα. Στο ένα τμήμα υπάρχει ένα μη επιβλεπόμενο δίκτυο, με σκοπό να πραγματοποιείται εξαγωγή των χαρακτηριστικών. Ενώ στο δεύτερο τμήμα υπάρχει ένα επιβλεπόμενο δίκτυο, με σκοπό να πραγματοποιείται η ταξινόμηση.
- Στη δεύτερη μορφή, η μηχανή έχει τη μορφή ενός δικτύου πρόσθιας τροφοδότησης, με τη χρήση ενός αλγορίθμου επιβλεπόμενης εκπαίδευσης. Η διαδικασία για την εξαγωγή των χαρακτηριστικών του δικτύου, στη συγκεκριμένη μορφή πραγματοποιείται στο κρυφό επίπεδο από τις υπολογιστικές μονάδες του.

### **2.4.3 Έλεγχος**

Μια ακόμα λειτουργία μάθησης αποτελεί ο έλεγχος μιας μονάδας. Ο έλεγχος είναι εφικτό να υλοποιηθεί με χρήση των νευρωνικών δικτύων. Με την έκφραση «έλεγχος μονάδας» αναφερόμαστε σε μια διεργασία ενός συστήματος που χρειάζεται η κατάσταση του να παραμένει ελεγχόμενη.

## **2.5 Αλγόριθμοι Μάθησης**

Όπως έχει ήδη αναφερθεί, υπάρχουν πολλοί και διαφορετικοί αλγόριθμοι μηχανικής μάθησης για τα τεχνητά νευρωνικά δίκτυα. Οι αλγόριθμοι μάθησης μπορούν να ταξινομηθούν σε διάφορες κατηγορίες με βάση το είδος του προβλήματος που καλούνται να διαχειριστούν.

Ως παράδειγμα αλγορίθμου της μηχανικής μάθησης μπορούν να αναφερθεί η μάθηση κατά Bayes. Πρόκειται για μια γνωστή και πολύ χρησιμοποιούμενη διαδικασία για τη δημιουργία ταξινομητών. Η συγκεκριμένη προσέγγιση έχει χρησιμοποιηθεί επιτυχώς από ένα πλήθος διαφορετικών εφαρμογών. Το αντικείμενο το οποίο υποβάλλεται σε διαδικασία μάθησης με έναν αλγόριθμο της κατηγορίας αυτής έχει άμεση σχέση με την κατανομή της πιθανότητας που διαθέτουν όλα τα στιγμιότυπα του εκάστοτε προβλήματος σύμφωνα με την κλάση τους. (N Ζώρης, 2014)

Οι διαδικασίες μάθησης με αλγορίθμους μάθησης κατά Bayes έχουν πολλά πλεονεκτήματα. Συγκεκριμένα, το μεγαλύτερο πλεονέκτημα που διαθέτουν είναι η ικανότητα αξιολόγησης των περιπτώσεων που μπορεί να καταλήξει ο συγκεκριμένος αλγόριθμος εκπαίδευσης. Ένα ακόμα σημαντικό πλεονέκτημα είναι η διάθεση ενός μέτρου σύγκρισης, εν αντιθέσει με άλλες μορφές μηχανικής μάθησης.

# ΚΕΦΑΛΑΙΟ 3 - ΝΕΥΡΩΝΑΣ PERCEPTRON

---

## 3.1 Εισαγωγή

Στα χρόνια κατά τα οποία ο τομέας των Νευρωνικών Δικτύων ξεκίνησε την ανάπτυξη και εξέλιξη του (1943-1958) δεν ήταν λίγοι οι ερευνητές όπου ξεχώρισαν με τα έργα τους πάνω στον συγκεκριμένο κλάδο. Στη συνέχεια, αναφερόμαστε σε ορισμένους επιστήμονες με τα σημαντικά τους επιτεύγματα. (Haykin, 2009)

- **McCulloch και Pitts**

Οι ίδιοι, το 1943 εισήγαγαν την ιδέα, ότι τα νευρωνικά δίκτυα μπορούν να αντιμετωπιστούν ως υπολογιστικές μηχανές.

- **Hebb**

Το 1949, ο Hebb εισήγαγε τον πρώτο κανόνα αυτό-οργανωμένης μάθησης (Self-organized Learning).

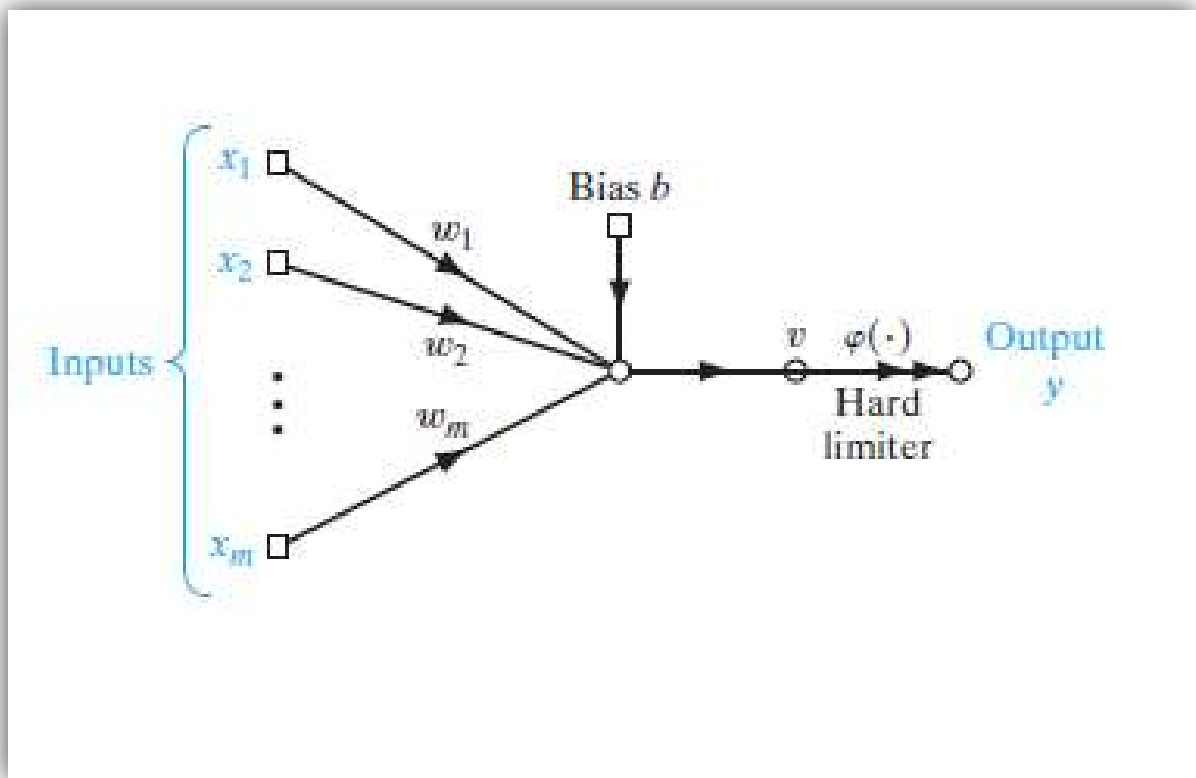
- **Rosenblatt**

Το 1958, ο Rosenblatt εισήγαγε την ιδέα του Perceptron. Ανέφερε το Perceptron ως το πρώτο πρότυπο για μάθηση με επίβλεψη.

## 3.2 Δίκτυο Perceptron

Με τον όρο «Perceptron» αναφερόμαστε στο μοντέλο αισθητήρα του Rosenblatt. Πρόκειται για την πιο απλή μορφή ενός Νευρωνικού Δικτύου (Haykin, 2009). Ο λόγος είναι ότι αποτελείται από έναν μόνο Νευρώνα. Αποτελεί το πρώτο μοντέλο Τεχνητού Νευρωνικού Δικτύου το οποίο δημιουργήθηκε και αποτέλεσε βάση για την μετέπειτα ανάπτυξη πιο περίπλοκων και σύνθετων δικτύων.

Στην Εικόνα 14 , παρουσιάζετε ο αισθητήρας Perceptron. Όπως είναι ευδιάκριτο και από το εικόνα που ακολουθεί το συγκεκριμένο δίκτυο αποτελείται αποκλειστικά από ένα πλήθος εισόδων ( $x_1, x_2 \dots x_n$ ) και μια παραγόμενη έξοδο  $y$ .



**Εικόνα 21 - Αισθητήρας Perceptron**

Στο δίκτυο Perceptron υπάρχουν μόνο οι συνδέσεις ανάμεσα στις εισόδους  $x_1, x_2 \dots x_n$  με τον αντίστοιχο νευρώνα. Όλες αυτές οι συνδέσεις διαθέτουν και από ένα αντίστοιχο συναπτικό βάρος  $w_i$ . Τα συναπτικά βάρη των συνδέσεων παρουσιάζουν την επίδραση του σήματος που εισέρχεται στον νευρώνα. Από μόνη της εξ ολοκλήρου η τιμή των βαρών δεν έχει τόση σημασία όση έχει το γινόμενο της εισόδου με το αντίστοιχο βάρος ( $x_i w_i$ ).

Ο συγκεκριμένος νευρώνας που αναπαρίστανται και στην παραπάνω εικόνα βασίζεται στο πρότυπο McCulloch – Pitts, το οποίο περιγράφηκε σε προηγούμενο κεφάλαιο. Για το νευρώνα αυτό υπάρχουν οι δύο συγκεκριμένες εξισώσεις.

Η πρώτη εξίσωση είναι η ακόλουθη:

$$u = \sum_{i=1}^m w_i x_i + b \quad (3.1)$$

Όσον αφορά την έξοδο του συγκεκριμένου νευρώνα γνωρίζουμε ότι είναι η  $y=f(u)$ . Η έξοδος  $y$  έχει την ακόλουθη εξίσωση, όπου αποτελεί και η δεύτερη συνάρτηση αυτού του νευρώνα.

$$y = f \left( \sum_{i=1}^m w_i x_i + b \right) \quad (3.2)$$

Στους παραπάνω τύπους με το σύμβολο « $u$ » αναφερόμαστε στην συνολική διέγερση του νευρώνα. Πιο συγκεκριμένα, το  $u$  συμβολίζει το ολικό σήμα που εισέρχεται σε αυτόν τον νευρώνα. Το σύμβολο « $f$ » είναι η συνάρτηση ενεργοποίησης. Με τη συγκεκριμένη συνάρτηση απεικονίζεται το σήμα εισόδου με τη μορφή διανύσματος ( $x = [x_1, x_2, \dots, x_n]^T$ ) στην αντίστοιχη έξοδο  $y$  του νευρώνα. (Haykin, 2009)

Παρατηρείται ότι η συνολική διέγερση  $u$  του συγκεκριμένου νευρώνα είναι θετική εάν η τιμή του αθροίσματος  $\sum_{i=1}^m w_i x_i$  υπερβεί το κατώφλι ενεργοποίησης.

Μέσω της συνάρτησης μεταφοράς του νευρώνα, η οποία τροφοδοτείται από το συνολικό σήμα  $u$  που λαμβάνει, παράγεται η έξοδος του. Πρόκειται για μια συνάρτηση μη γραμμική.

Για την περίπτωση του δικτύου Perceptron διακρίνονται δύο διαφορετικές μορφές για τη συνάρτηση κατωφλιού, οι οποίες περιγράφονται παρακάτω.



1.  $f(u) = \begin{cases} 1, & \text{αν το } u \text{ είναι θετικό } (u > 0) \\ 0, & \text{αν το } u \text{ είναι αρνητικό ή μηδενικό } (u \leq 0) \end{cases}$
  
2.  $f(u) = \begin{cases} 1, & \text{αν το } u \text{ είναι θετικό } (u > 0) \\ -1, & \text{αν το } u \text{ είναι αρνητικό ή μηδενικό } (u \leq 0) \end{cases}$

Από τα όσα αναφέρθηκαν παραπάνω, φτάνουμε στο συμπέρασμα ότι η διαδικασία ενεργοποίησης του συγκεκριμένου νευρώνα εξαρτάται από τρεις παραμέτρους. Πιο συγκεκριμένα, οι παράμετροι αυτοί είναι οι τιμές που δέχεται ως εισόδους, οι τιμές των βαρών των συνδέσεων του και τέλος η τιμή που έχει το κατώφλι. (Haykin, 2009)

Το κατώφλι  $\theta$  μπορεί να ερμηνευτεί και ως ένα ακόμα εσωτερικό συναπτικό βάρος. Πρόκειται για το  $w_0 = \theta$ , το οποίο αντιστοιχίζεται με μια είσοδο όπου ισχύει  $x_0 = -1$ . Η δράση του συγκεκριμένου συναπτικού βάρους είναι ίδια με τα υπόλοιπα συναπτικά βάρη του Perceptron. (Αργυράκης, 2001) Έτσι, η μορφή της  $u$  μπορεί να είναι η ακόλουθη:

$$u = \sum_{i=1}^n w_i x_i - \theta = \sum_{i=1}^n w_i x_i + w_0 x_0 = \sum_{i=0}^n w_i x_i$$

Η παραπάνω εξίσωση περιλαμβάνει το εσωτερικό γινόμενο των βαρών των συνδέσεων  $w$  και των εισόδων  $x$  του δικτύου. Με τον όρο εσωτερικό γινόμενο αναφερόμαστε στο  $u = w^T x$ . Στο διάνυσμα εισόδου και στο διάνυσμα των βαρών προστίθεται από μια επιπλέον τιμή. Πιο συγκεκριμένα, στο διάνυσμα εισόδου προστίθεται η τιμή  $x_0$ , η οποία είναι ίση με  $-1$  και παραμένει σταθερή. Η μορφή του διανύσματος εισόδου μετά από τη συγκεκριμένη προσθήκη έχει την εξής μορφή:

$$[x_0, x_1, x_2, \dots, x_n]^T$$

Η μορφή του διανύσματος των βαρών ύστερα από τη συγκεκριμένη προσθήκη έχει τη μορφή:

$$[w_0, w_1, w_2, \dots, w_n]^T$$

### 3.3 Ταξινόμηση προτύπων

Ο κύριος στόχος για τη δημιουργία του δικτύου Perceptron είναι να χρησιμοποιείται κατάλληλα προκειμένου να επιτυγχάνεται η ταξινόμηση των προτύπων. Προκειμένου να πετύχει τη σωστή λειτουργία του, θα πρέπει να χρησιμοποιείται ο λεγόμενος *γραμμικός διαχωρισμός*.

Ο Rosenblatt δημιούργησε έναν αλγόριθμο, ο οποίος μπορεί να εκπαιδεύει το συγκεκριμένο Νευρωνικό Δίκτυο προσαρμόζοντας κατάλληλα τις παραμέτρους του. Ο ίδιος ύστερα από τις μελέτες του κατέληξε στο εξής συμπέρασμα. Εάν τα διανύσματα που χρησιμοποιούνται για την εκπαίδευση του αισθητήρα Perceptron προκύπτουν από γραμμικά διαχωρίσιμες κλάσεις, τότε προκύπτει ως αποτέλεσμα η σύγκλιση του αλγορίθμου Perceptron σε πεπερασμένο πλήθος επαναλήψεων. Ως συνέπεια της σύγκλισης αυτής, τοποθετείται η διαχωριστική επιφάνεια απόφασης ανάμεσα στις δύο κλάσεις υπό την μορφή υπερεπιπέδου.

Η απλή μορφή του Perceptron, ο οποίος αποτελείται από έναν μόνο νευρώνα, στοχεύει στην διαδικασία της ταξινόμησης των προτύπων ανάμεσα σε μόνο δύο κλάσεις. Πιο συγκεκριμένα, στοχεύει στην πλέον κατάλληλη και σωστή ταξινόμηση των εισόδων  $x_1, \dots, x_n$  σε μια από τις δύο αυτές κλάσεις ( $C_1, C_2$ ).

Υπάρχει ένας συγκεκριμένος κανόνας στον οποίο βασίζεται η διαδικασία της απόφασης όσον αφορά την ταξινόμηση των προτύπων. Ο κανόνας αυτός αναφέρει ότι αν η έξοδος του νευρώνα είναι  $y = +1$  τότε στην πρώτη κλάση  $C_1$  ανατίθεται το συγκεκριμένο σημείο το οποίο δείχνουν οι είσοδοι του. Αντίθετα, αν η έξοδος του έχει την τιμή  $y = -1$ , τότε το σημείο αυτό ανατίθεται στην δεύτερη κλάση  $C_2$ .

Για την πλήρη και σωστή κατανόηση του Perceptron είναι απαραίτητο να αναφερθεί ότι η σχεδίαση των περιοχών απόφασης πραγματοποιείται στον n-

διάστατο χώρο για τις  $n$  μεταβλητές εισόδου. Σε αυτή την περίπτωση του απλού Perceptron που αποτελείται από έναν και μόνο νευρώνα, διακρίνονται δύο περιοχές απόφασης. (Rojas, 1996) Οι συγκεκριμένες περιοχές απόφασης χωρίζονται από ένα υπερεπίπεδο. Το υπερεπίπεδο αυτό μπορεί να οριστεί με τους ακόλουθους τρόπους:

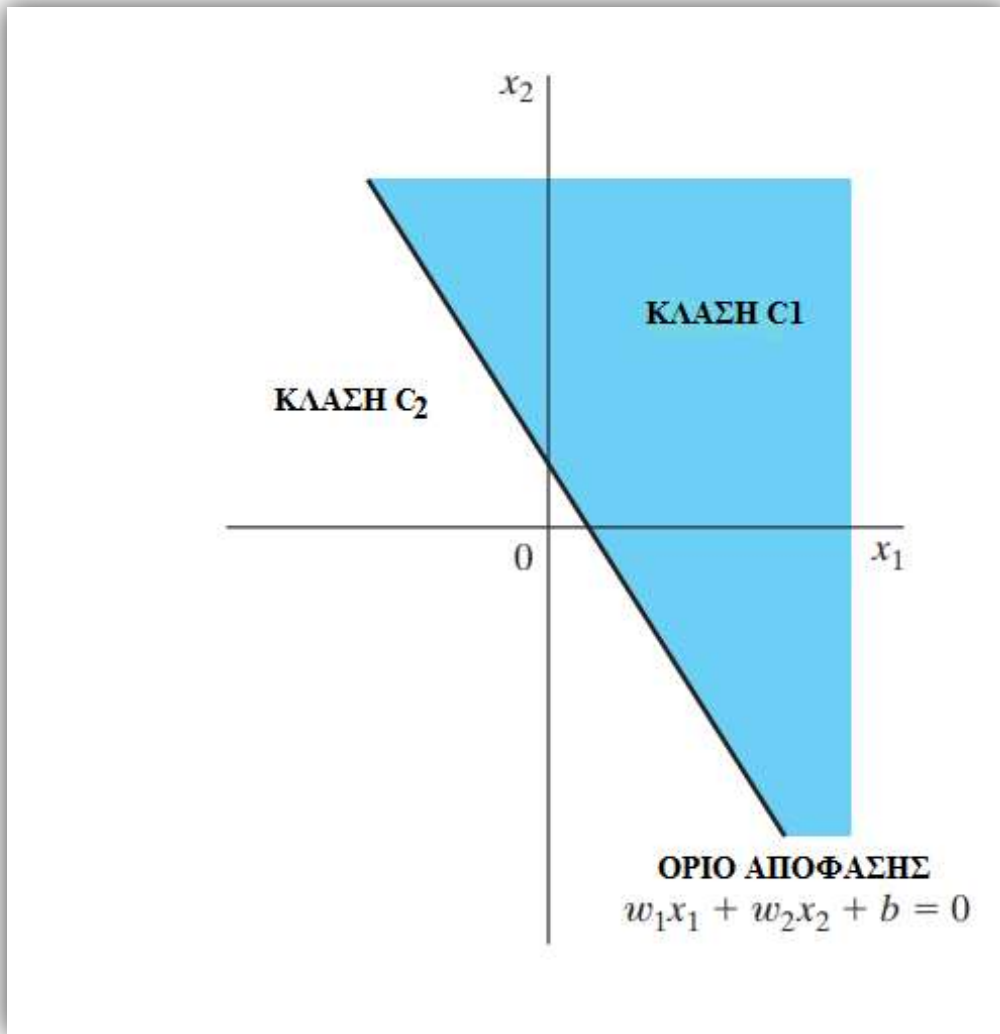
$$1. u = \sum_{i=1}^m w_i x_i - \theta = 0$$

$$2. \sum_{i=1}^m w_i x_i + b = 0$$

Ως παράδειγμα κατανόησης μπορούμε να αναφέρουμε την περίπτωση όπου υπάρχουν μόνο δύο είσοδοι ( $x_1, x_2$ ). Στη συγκεκριμένη περίπτωση, το υπερεπίπεδο αναπαρίστανται με μια ευθεία γραμμή. Η γραμμή αυτή χρησιμεύει ως ένα διαχωριστικό όριο για την απόφαση ανάμεσα στις κλάσεις  $C_1$  και  $C_2$  (Αργυράκης, 2001). Η εξίσωση που αναφέρεται σε αυτό το όριο απόφασης είναι η εξής:

$$w_1x_1 + w_2x_2 + b = 0$$

Στην Εικόνα 15 που ακολουθεί αναπαριστάνεται το υπερεπίπεδο για τις δύο μεταβλητές εισόδου  $x_1, x_2$ . (Haykin, 2009)



**Εικόνα 22 - Αναπαράσταση του υπερεπιπέδου για την ταξινόμηση δύο κλάσεων**

Τα σημεία που υπάρχουν πάνω από την ευθεία γραμμή, στην μπλε επιφάνεια όπως φαίνεται από την παραπάνω εικόνα αποτελούν σημεία  $(x_1, x_2)$ , τα οποία έχουν την ιδιότητα  $u > 0$ . Τα σημεία αυτά βρίσκονται στην πρώτη κλάση ( $C_1$ ), διότι η έξοδος του συγκεκριμένου νευρώνα είχε την τιμή  $+1$  ( $y=+1$ ). Τα σημεία με τιμή  $u < 0$  είναι τα σημεία που βρίσκονται κάτω από την ευθεία γραμμή, στην λευκή επιφάνεια. Τα συγκεκριμένα σημεία ανήκουν στην δεύτερη κλάση, διότι η έξοδος του νευρώνα είχε την τιμή  $-1$  ( $y=-1$ ). Από την άλλη τα σημεία για τα οποία ισχύει  $u = 0$ , είναι τα σημεία που βρίσκονται ακριβώς πάνω στη συγκεκριμένη ευθεία.

Το  $b$  είναι η πόλωση και αναφέρεται συχνά και με τον όρο *bias*. Η τιμή της πόλωσης μετακινεί από την αρχή των αξόνων το όριο απόφασης. (Widrow, 1998)

### 3.4 Η εκπαίδευση του αισθητήρα

Η εκπαίδευση του Νευρωνικού Δικτύου Perceptron είναι πολύ σημαντική ώστε να μπορεί να επιτύχει το σκοπό της ύπαρξής του. Ο στόχος του είναι να βρεθεί η διαχωριστική γραμμή για τις κλάσεις  $C_1$  και  $C_2$ . Αυτό επιτυγχάνεται με την αυτόματη εκπαίδευση των παραμέτρων του.

Στόχος είναι η εύρεση ενός διανύσματος από βάρη  $w$  για το οποίο να ισχύει μια από τις ακόλουθες ανισότητες:

$$1. w^T x > 0,$$

για τις εισόδους που βρίσκονται στην  $C_1$  κλάση.

$$2. w^T x \leq 0,$$

για τις εισόδους που βρίσκονται στην  $C_2$  κλάση.

Για τη διαδικασία μάθησης του Νευρωνικού Δικτύου Perceptron χρησιμοποιεί η εκπαίδευση με επίβλεψη. Πιο συγκεκριμένα, στο δίκτυο αυτό παρέχεται η επιθυμητή έξοδος  $d$  για κάθε είσοδο εκπαίδευσης  $x$  που δέχεται με τη χρήση ενός αλγορίθμου επανάληψης. (Jantzen, 1998)

Ως βασικό κανόνα μάθησης χρησιμοποιείται ο κανόνας fixed increment - σταθερής αύξησης (WIDROW, 1990). Σύμφωνα με τον κανόνα αυτό όλα τα πρότυπα εισέρχονται στο συγκεκριμένο δίκτυο με κυκλικό τρόπο. Όταν το πλήθος των εισερχόμενων προτύπων τελειώσει, η διαδικασία επαναλαμβάνεται από την αρχή. Το διάνυσμα των βαρών αλλάζει μόνο όταν παρατηρηθεί ένα σφάλμα ταξινόμησης. Αυτό συμβαίνει όταν η έξοδος για ένα πρότυπο είναι διαφορετική από την επιθυμητή έξοδο.

### 3.5 Αλγόριθμος σύγκλισης

Στη συνέχεια γίνεται μια περιγραφή του αλγορίθμου σύγκλισης του Νευρωνικού Δικτύου Perceptron.

Ως διάνυσμα εισόδου θεωρούμε το εξής:

$$x(n) = [+1, x_1(n), \dots, x_n(n)]^T$$

Ως διάνυσμα των βαρών θεωρούμε το ακόλουθο:

$$w(n) = [b, w_1(n), \dots, w_n(n)]^T$$

Ο αλγόριθμος σύγκλισης ακολουθεί συγκεκριμένα βήματα για την υλοποίησή του. Τα βήματα αυτά αναλύονται στη συνέχεια.

1. Στο πρώτο βήμα του αλγορίθμου υπάρχει η αρχικοποίηση. Με τον όρο αρχικοποίηση αναφερόμαστε στην απόδοση της μηδενικής τιμής στο διάνυσμα που περιέχει τα βάρη. Πιο συγκεκριμένα, έχουμε  $w(0)=0$  και  $n=1$ .
2. Το επόμενο βήμα του αλγορίθμου είναι η ενεργοποίηση. Όταν φτάσει η επανάληψη  $n$ , έχει ως αποτέλεσμα την ενεργοποίηση του δικτύου Perceptron. Στο σημείο αυτό στο δίκτυο χρησιμοποιείται το διάνυσμα εισόδου  $x(n)$ .
3. Στο τρίτο βήμα κατά σειρά, πραγματοποιείται ο υπολογισμός της πραγματικής εξόδου του δικτύου Perceptron. Ο υπολογισμός αυτός υλοποιείται με τους ακόλουθους τύπους:

$$u(n) = w^T(n) x(n)$$

$$y(n) = \text{sgn}(u(n)) = \text{sgn}[w^T(n) x(n)]$$

Με τον συμβολισμό  $\text{sgn}$  που υπάρχει στους παραπάνω τύπους αναφερόμαστε στη συνάρτηση προσήμου, η οποία υπολογίζεται σύμφωνα με τον παρακάτω τύπο:

$$\text{sgn}(u) = \begin{cases} +1 & \text{αν } u > 0 \\ -1 & \text{αν } u < 0 \end{cases}$$

4. Στο επόμενο βήμα του αλγορίθμου πραγματοποιείται η μεταβολή του διανύσματος βαρών του δικτύου. Η προσαρμογή αυτή υλοποιείται σύμφωνα με τον κανόνα που ακολουθεί:

$$w(n+1) = w(n) + \eta [d(n) - y(n)] x(n)$$

Η μεταβλητή  $\eta$  που υπάρχει στον παραπάνω κανόνα αναφέρεται στο ρυθμό εκπαίδευσης. Πρόκειται ουσιαστικά για το βήμα που καθορίζει το μέγεθος της διόρθωσης. Οι τιμές που μπορεί να πάρει η συγκεκριμένη μεταβλητή ανήκουν στο διάστημα  $[0 \ 1]$ .

Επίσης, η μεταβλητή  $d(n)$  που υπάρχει και αυτή στον παραπάνω κανόνα υπολογίζεται από τον εξής τύπο:

$$d(n) = \begin{cases} +1, & \text{αν } X(n) \text{ ανήκει στην κλάση } C1 \\ -1, & \text{αν } X(n) \text{ ανήκει στην κλάση } C2 \end{cases}$$

5. Στο βήμα αυτό πραγματοποιείται αύξηση κατά μια μονάδα του βήματος  $n$ . Στο σημείο αυτό ο αλγόριθμος συνεχίζει, επιστρέφοντας στο βήμα 2.

Η διαδικασία με τα πέντε βήματα που αναλύθηκε προηγουμένως αποτελεί τον αλγόριθμο σύγκλισης για το Νευρωνικό Δίκτυο Perceptron. Η διαδικασία αυτή πραγματοποιείται διαρκώς. Αν ένα πρότυπο ταξινομήθηκε λάθος στη μια επανάληψη που έγινε, στην επόμενη επανάληψη ή θα ταξινομηθεί στο σωστό σημείο ή θα φτάσει πιο κοντά στη σωστή, ταξινομημένη θέση του.

### 3.6 Μειονέκτημα του δικτύου Perceptron

Υπάρχει ένα πολύ σημαντικό μειονέκτημα που παρατηρείται σε αυτό το δίκτυο Perceptron. Αναφερόμαστε στο γεγονός ότι δε μπορούν με το συγκεκριμένο δίκτυο να επιλυθούν μη γραμμικά διαχωρίσιμα προβλήματα. Όλα σχεδόν τα προβλήματα που υπάρχουν στην πραγματικότητα είναι μη γραμμικά διαχωρίσιμα. Υπάρχει πολύ μικρός αριθμός από γραμμικά διαχωρίσιμα προβλήματα στον πραγματικό κόσμο. Τα προβλήματα αυτά όμως έχουν επιλυθεί με πιο απλές και εύκολες μεθόδους. (N Ζώρης, 2014)

Για να μπορέσουν οι επιστήμονες του συγκεκριμένου κλάδου να εξαλείψουν αυτό το πολύ σημαντικό μειονέκτημα που υπάρχει στο δίκτυο Perceptron με έναν μόνο νευρώνα ξεκίνησαν να μελετούν και να αναπτύσσουν το δίκτυο Perceptron που αποτελείται από πολλά επίπεδα. Το συγκεκριμένο δίκτυο Perceptron αποτελεί ένα ιδιαίτερα περίπλοκο δίκτυο, καθώς αποτελείται από μεγαλύτερο πλήθος νευρώνων.

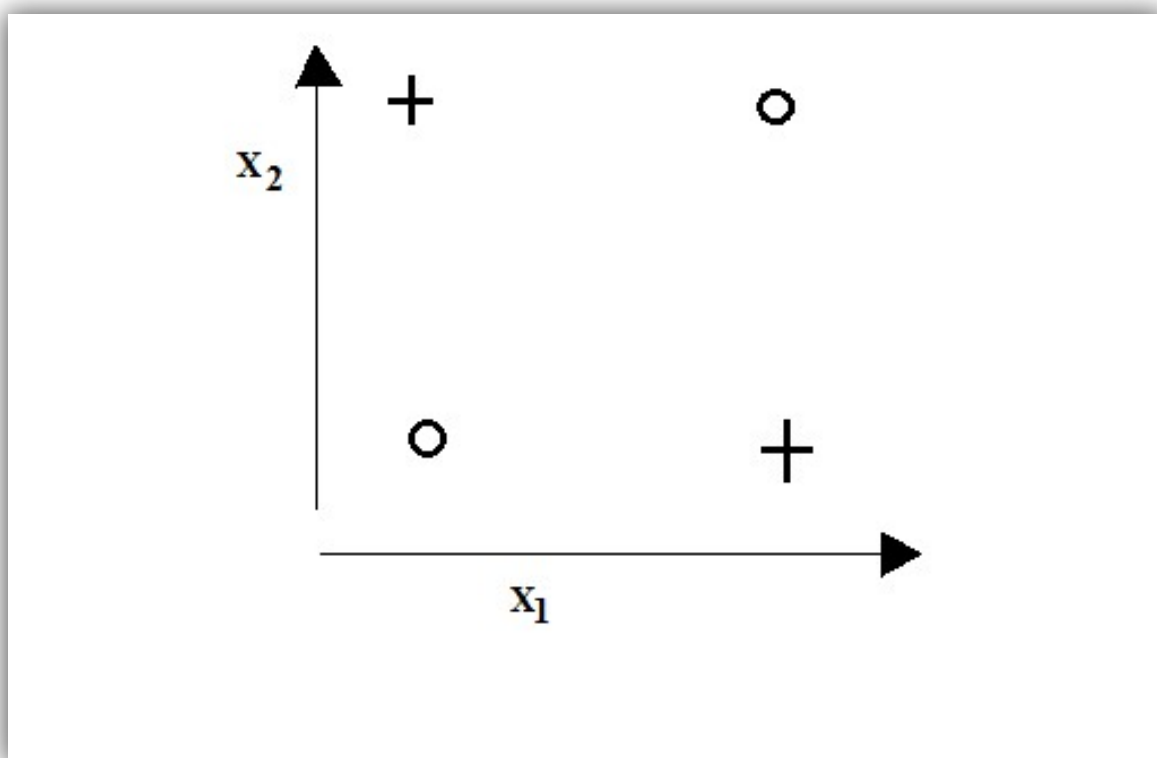
### 3.7 Πρόβλημα Συνάρτησης XOR

Ως παράδειγμα μη γραμμικού διαχωρίσιμου προβλήματος μπορούμε να αναφερθούμε στην συνάρτηση XOR. Η συνάρτηση XOR αποτελεί ένα σχετικά



απλό παράδειγμα του είδους αυτού, και είναι αδύνατη η αναπαράσταση της με δίκτυο Perceptron ενός νευρώνα, όπως δείχνει και η Εικόνα 16 που βρίσκεται παρακάτω.

Το 1961, οι Minsky και Selfridge άσκησαν την πρώτη κριτική όσον αφορά το δίκτυο Perceptron ενός επιπέδου, καθώς επισήμαναν την αδυναμία υλοποίησης της συνάρτησης ισοτιμίας 2 bits, δηλαδή τη συνάρτηση XOR. (GALLANT, 1990)



**Εικόνα 23 - Πρόβλημα XOR**

Στο πρόβλημα της συνάρτησης XOR υπάρχουν δύο κλάσεις. Η κάθε κλάση αποτελείται από δύο μόνο πρότυπα. Στην πρώτη κλάση τα πρότυπα συμβολίζονται με το σύμβολο + και πρόκειται για τα διανύσματα εισόδων  $[1,0]$ ,  $[0,1]$ . Ενώ αντιθέτως, στη δεύτερη κλάση τα πρότυπα έχουν το σύμβολο o και πρόκειται για τις εισόδους  $[0,0]$ ,  $[1,1]$ . (GALLANT, 1990)

Για να μπορέσει η συνάρτηση XOR να πραγματοποιηθεί σωστά θα πρέπει για την πρώτη κλάση να έχουμε έξοδο  $y=1$  και για τη δεύτερη κλάση να έχουμε την έξοδο  $y=0$ . Όπως απεικονίζεται και στην Εικόνα 16, δεν μπορεί να υπάρξει μια ευθεία γραμμή, τέτοια ώστε να διαχωρίζει σωστά τις συγκεκριμένες κλάσεις.

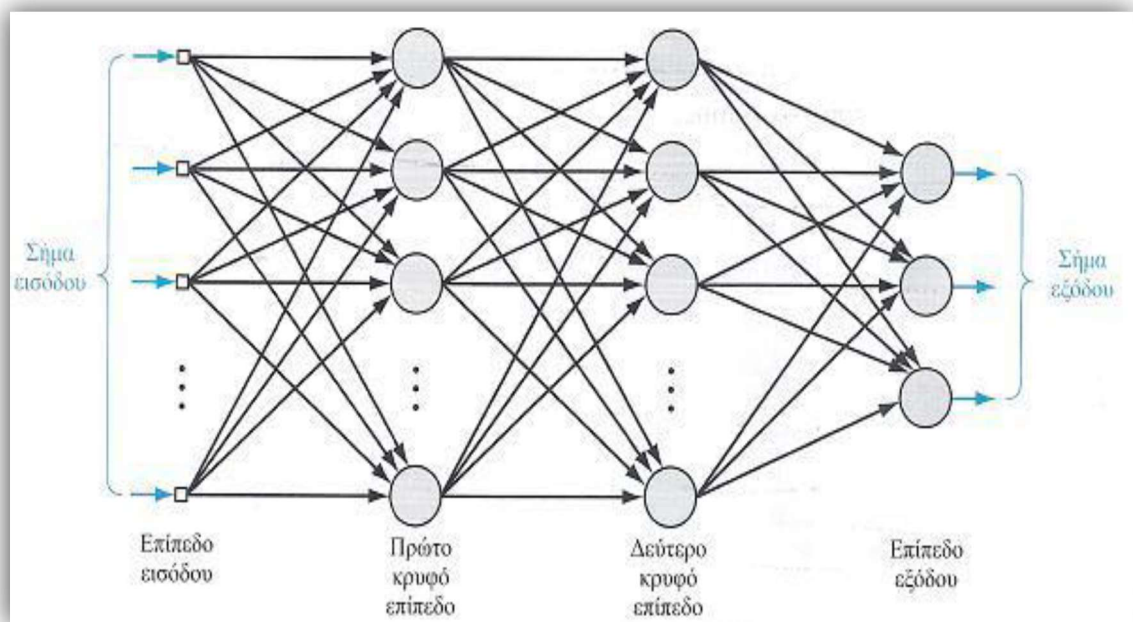
Καταλήγουμε στο συμπέρασμα ότι το συγκεκριμένο πρόβλημα της συνάρτησης XOR δεν μπορεί να επιλυθεί με το νευρωνικό δίκτυο Perceptron με μόνο έναν νευρώνα. (GALLANT, 1990)

# ΚΕΦΑΛΑΙΟ 4 – PERCEPTRON ΠΟΛΛΑΠΛΩΝ ΕΠΙΠΕΔΩΝ

## 4.1 Πολυεπίπεδο δίκτυο Perceptron – MLP

Όπως έχει ήδη αναφερθεί και στο προηγούμενο κεφάλαιο, η πιο απλή μορφή τεχνητού νευρωνικού δικτύου αποτελεί το δίκτυο Perceptron, το οποίο αποτελείται από έναν και μόνο νευρώνα. Το συγκεκριμένο δίκτυο δεν μπορεί να χρησιμοποιηθεί προκειμένου να επιλυθούν αρκετά από τα προβλήματα που καλείται να επιλύσει. Για το λόγο αυτό οι επιστήμονες του κλάδου των Νευρωνικών Δικτύων εξέλιξαν τον απλό Perceptron ενός νευρώνα σε ένα δίκτυο Perceptron πολλών επιπέδων (Multi-Layer Perceptron - MLP). Πρόκειται ουσιαστικά για μια γενίκευση του απλού δικτύου Perceptron. (Haykin, 2009)

Στην Εικόνα 17 που ακολουθεί παρουσιάζεται η αρχιτεκτονική ενός πολυεπίπεδου δικτύου Perceptron. (Haykin, 2009)



Εικόνα 24 - Πολυεπίπεδο δίκτυο Perceptron

Όπως είναι φανερό και από την παραπάνω εικόνα, ένα πολυεπίπεδο δίκτυο Perceptron απαρτίζεται από πολλά επίπεδα. Συγκεκριμένα, ένα τέτοιο δίκτυο διαθέτει ένα πλήθος αισθητήρων που υπάρχουν στο επίπεδο της εισόδου καθώς και ένα πλήθος υπολογιστικών κόμβων που αποτελούν το επίπεδο της εξόδου. Ανάμεσα από τα δύο αυτά επίπεδα υπάρχουν ένα ή και περισσότερα κρυφά επίπεδα (hidden layers). (Haykin, 2009) Στην εικόνα 17, το δίκτυο που απεικονίζεται είναι ένα δίκτυο τριών επιπέδων που αποτελείται από δύο κρυφά επίπεδα.

Όλοι οι νευρώνες του πολυεπίπεδου δικτύου Perceptron που βρίσκονται σε οποιοδήποτε στρώμα του, έστω  $\ell$  διαθέτουν την ιδιότητα να τροφοδοτούν του νευρώνες που βρίσκονται στο αμέσως επόμενο επίπεδο  $\ell + 1$  και μόνο αυτούς. Οι νευρώνες του επιπέδου  $\ell$  επίσης τροφοδοτούνται από τους νευρώνες του ακριβώς προηγούμενου επιπέδου  $\ell - 1$ . Αυτό αποτελεί και το βασικό χαρακτηριστικό των πολυεπίπεδων δικτύων MLP. (Shepherd, 1997)

Τέτοιου είδους νευρωνικά δίκτυα πολλών επιπέδων, όπως είναι και το νευρωνικό δίκτυο της Εικόνας 17 καλούνται δίκτυα **πλήρως διασυνδεδεμένα**. Πιο αναλυτικά, σε ένα πλήρως διασυνδεδεμένο δίκτυο κάθε νευρώνας που βρίσκεται σε ένα επίπεδο συνδέεται με όλους τους νευρώνες που υπάρχουν στο ακριβώς προηγούμενο επίπεδο του. Τα σήματα που εισέρχονται μέσα στο συγκεκριμένο δίκτυο κινούνται προς μια μόνο κατεύθυνση. Συγκεκριμένα, τα σήματα διαπερνούν το δίκτυο από τα αριστερά προς τα δεξιά.

Κάθε νευρωνικό δίκτυο πολλών επιπέδων διαθέτει κάποια συγκεκριμένα χαρακτηριστικά. Τα χαρακτηριστικά αυτά αναφέρονται και αναλύονται στη συνέχεια:

- Όλοι οι νευρώνες που υπάρχουν και απαρτίζουν το νευρωνικό δίκτυο διαθέτουν μια μη γραμμική συνάρτηση ενεργοποίησης. Η σημαντική ιδιότητα που διαθέτει η συνάρτηση ενεργοποίησης είναι το γεγονός ότι είναι σε όλα τα σημεία παραγωγίσιμη. Για τη συνάρτηση αυτή χρησιμοποιείται τις περισσότερες φορές ως μη γραμμική συνάρτηση η σιγμοειδή συνάρτηση. Η σιγμοειδή συνάρτηση ενεργοποίησης έχει τον ακόλουθο τύπο:

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

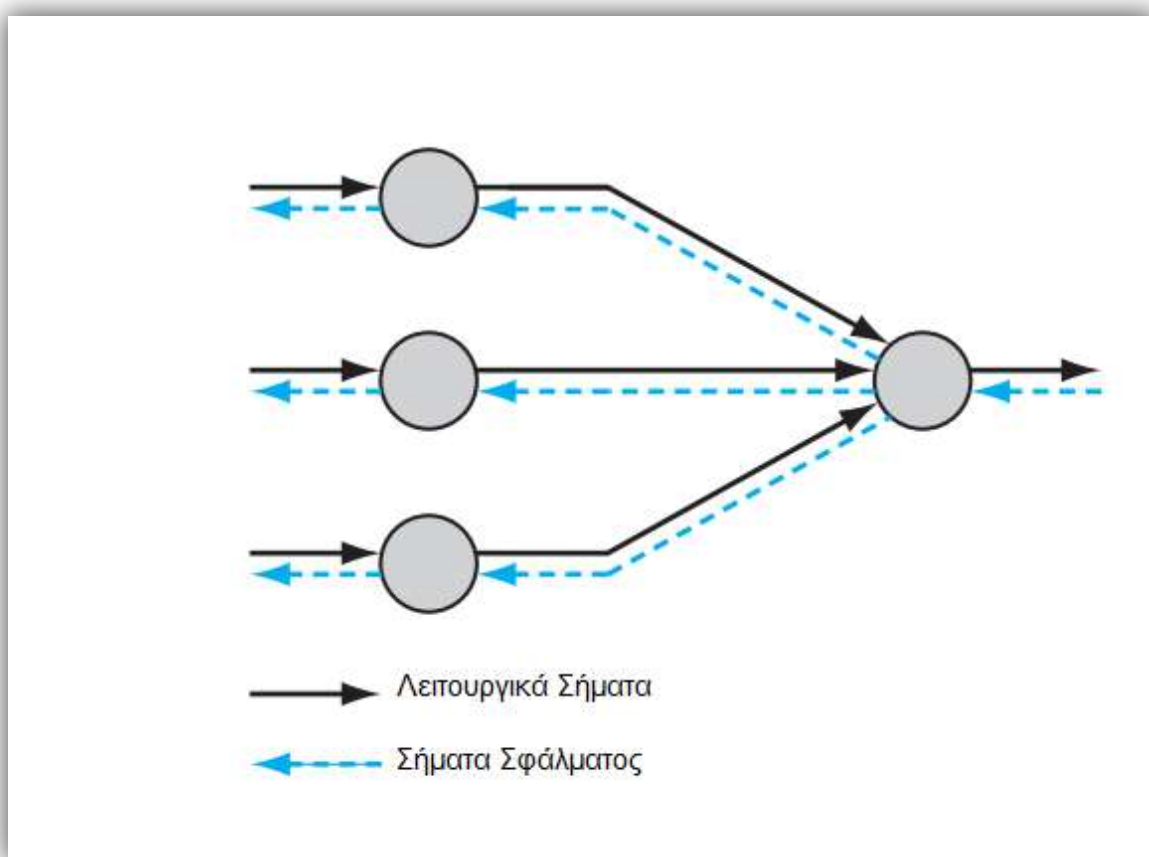
- Η παραπάνω συνάρτηση ενεργοποίησης διαθέτει συνεχείς τιμές. Εν αντιθέσει με τη βηματική συνάρτηση ενεργοποίησης, που περιγράφηκε σε προηγούμενο κεφάλαιο, η οποία διαθέτει μόνο δυαδικές τιμές. Επίσης, η βηματική συνάρτηση ενεργοποίησης αποτελεί μια μη παραγωγίσιμη συνάρτηση.

Είναι ιδιαίτερα σημαντικό το γεγονός ότι η συνάρτηση ενεργοποίησης των νευρώνων του νευρωνικού δικτύου πρέπει να είναι μια παραγωγίσιμη συνάρτηση. Ο λόγος για τον οποίο είναι τόσο σημαντική αυτή η ιδιότητα είναι ότι σχεδόν όλοι οι κανόνες μάθησης που χρησιμοποιούνται για την εκπαίδευση του νευρωνικού δικτύου, χειρίζονται μεθόδους για την βελτιστοποίηση που περιλαμβάνουν διάφορες παραγώγους. (Wilde, 1996)

- Μια επιπλέον σημαντική ιδιότητα που διαθέτουν τα νευρωνικά δίκτυα πολλών επιπέδων είναι η μη γραμμικότητα. Αποτελεί μια απαραίτητη ιδιότητα για τα συγκεκριμένα δίκτυα. Ο λόγος είναι διότι αν υπήρχε γραμμικότητα στα δίκτυα πολλών επιπέδων, η σχέση εισόδου – εξόδου θα μειωνόταν στη μορφή ενός απλού δικτύου Perceptron που διαθέτει ένα και μόνο επίπεδο. (Αργυράκης, 2001)
- Ένα νευρωνικό δίκτυο πολλών επιπέδων μπορεί να περιέχει από ένα έως και περισσότερα κρυφά επίπεδα. Τα επίπεδα αυτά παραμένουν κρυφά τόσο για το επίπεδο εισόδου όσο και για το επίπεδο εξόδου. Ο σκοπός για την ύπαρξη των κρυφών επιπέδων σε ένα δίκτυο είναι η ικανότητα τους να παρέχουν γνώση για πολύπλοκες διεργασίες, δεδομένου ότι προσφέρουν τις σημαντικότερες ιδιότητες και χαρακτηριστικά από τις εισόδους του.
- Κάθε δίκτυο Perceptron που αποτελείται από πολλά επίπεδα διαθέτει επιπλέον τη δυνατότητα ενός υψηλού βαθμού διασύνδεσης. Τα συναπτικά βάρη των συνδέσεων συμβάλλουν σε αυτή τη ιδιότητα του δικτύου.
- Στα νευρωνικά δίκτυα πολλών επιπέδων MLP διακρίνονται δύο διαφορετικές κατηγορίες σημάτων. Στην πρώτη κατηγορία ανήκουν τα λειτουργικά σήματα (function signals), ενώ στη δεύτερη κατηγορία ανήκουν τα σήματα σφάλματος (error signals). (Haykin, 2009)

Τα λειτουργικά σήματα αποτελούν σήματα εισόδου του νευρωνικού δικτύου. Πρόκειται για σήματα τα οποία εισέρχονται στο δίκτυο MLP και στη συνέχεια διαπερνάνε το δίκτυο αυτό με φορά προς τα εμπρός. Καταλήγουν στην έξοδο του δικτύου. Στο σημείο αυτό εμφανίζονται ως σήματα εξόδου και εκτελούν μια σημαντική λειτουργία. Κατά το πέρασμα των λειτουργικών σημάτων από τους νευρώνες του νευρωνικού δικτύου πολλών επιπέδων υπολογίζονται από τα σήματα εισόδου του κάθε νευρώνα συναρτήσει των συναπτικών βαρών τους. (Haykin, 2009)

Τα σήματα σφάλματος (error signals) σχηματίζονται στους νευρώνες εξόδου ενός δικτύου Perceptron πολλών επιπέδων. Πρόκειται για τα σήματα τα οποία διαδίδονται με φορά προς τα πίσω, διασχίζοντας τα επίπεδα του κάθε δικτύου αυτού του τύπου.



**Εικόνα 25 - Απεικόνιση των σημάτων ενός MLP δικτύου**

Στην Εικόνα 18 που προηγείται, απεικονίζεται η κατεύθυνση που έχουν τόσο τα λειτουργικά σήματα όσο και τα σήματα σφάλματος που υπάρχουν σε ένα δίκτυο Perceptron πολλών επιπέδων.

Όπως έχει ήδη αναφερθεί, ένα δίκτυο Perceptron πολλών επιπέδων διαθέτει ένα ή και περισσότερα κρυφά επίπεδα. Κάθε ένας από τους κρυφούς νευρώνες του δικτύου υπολογίζει ένα σήμα και συγκεκριμένα το λειτουργικό σήμα του το οποίο θα εμφανιστεί στη συνέχεια στην έξοδο του. Στο σημείο αυτό πραγματοποιείται μια εκτίμηση για το διάνυμα κλίσης, με σκοπό να χρησιμοποιηθεί στη συνέχεια για την διάδοση του σήματος προς τα πίσω. (Noriega, 2005)

Τα προβλήματα τα οποία δεν μπορούν να υλοποιηθούν με τη χρήση ενός απλού Perceptron ενός επιπέδου, μπορούν να επιλυθούν με τη χρήση ενός Perceptron MLP πολλών επιπέδων.

Επίσης, ένα MLP δίκτυο μπορεί να χρησιμοποιήσει και να υλοποιήσει σωστά ορισμένες συναρτήσεις οι οποίες όμως δεν μπορούν να υλοποιηθούν σωστά από έναν απλό Perceptron. Όταν από ένα MLP δίκτυο γίνεται χρήση μιας σιγμοειδή συνάρτησης το δίκτυο αυτό έχει τη δυνατότητα να προσεγγίσει όσο το δυνατόν καλύτερα τις συναρτήσεις, οι οποίες είναι συνεχής. Οι συγκεκριμένες προσεγγίσεις που περιγράφηκαν καλούνται επίσης και καθολικές προσεγγίσεις – universal approximations. (Hassan Ramchoun, 2016)

Προκειμένου να γίνει απόλυτα κατανοητή η λειτουργία των νευρωνικών δικτύων που αποτελούνται από πολλά επίπεδα MLP, είναι σημαντικό να αναλυθεί ένα Perceptron δύο επιπέδων.

Πιο συγκεκριμένα, θεωρούμε ένα Νευρωνικό δίκτυο με δύο μόνο επίπεδα, το οποίο έχει συνολικά  $n$  εισόδους που είναι οι  $x_1, x_2, \dots, x_n$ . Θεωρούμε επίσης ότι στο πρώτο κρυφό επίπεδο του δικτύου αυτού υπάρχουν συνολικά  $p$  νευρώνες. Οι συγκεκριμένοι νευρώνες διαθέτουν την ιδιότητα της μη γραμμικότητας και επίσης έχουν ως συνάρτηση ενεργοποίησης την σιγμοειδή συνάρτηση  $f$ . Ο μαθηματικός τύπος που δίνει την έξοδο κάθε νευρώνα που βρίσκεται στο πρώτο κρυφό επίπεδο του νευρωνικού δικτύου είναι ο ακόλουθος: (N Ζώρης, 2014)

$$y_i = f \left( \sum_{j=1}^n w_{ij} x_j + \theta_i \right)$$

Όσον αφορά το δεύτερο επίπεδο του συγκεκριμένου Νευρωνικού δικτύου γνωρίζουμε ότι περιλαμβάνει έναν νευρώνα, ο οποίος είναι γραμμικός. Ο νευρώνας αυτός, υλοποιεί την λειτουργία της πρόσθεσης όλων των ενεργοποιήσεων του προηγούμενου επιπέδου του με τη συνεισφορά των συναπτικών τους βαρών  $\omega_1, \omega_2, \dots, \omega_p$ .

Σύμφωνα με τα παραπάνω που αναφέρθηκαν, η έξοδος του νευρωνικού αυτού δικτύου είναι η εξής:

$$g(x_1, x_2, \dots, x_n) = \sum_{i=1}^p \omega_i y_i - \theta$$

Συμπεραίνεται ότι ο αριθμός  $p$  ο οποίος είναι ένας ακέραιος αριθμός και οι τιμές των βαρών  $\omega_i, w_{ij}$  καθώς και  $\theta_i$  των νευρώνων του δικτύου συνεισφέρουν έτσι ώστε η τελική έξοδος του δικτύου αυτού  $g(x_1, x_2, \dots, x_n)$  να μπορεί να προσεγγίζει όσο το δυνατόν καλύτερα όλες τις συναρτήσεις  $f(x_1, x_2, \dots, x_n)$  που είναι συνεχής. Με τη σημαντική προϋπόθεση ότι τα στοιχεία  $x_i$  βρίσκονται στο διάστημα  $[0,1]$ .

Όσα αναφέρθηκαν προηγουμένως, αποτελούν μια πολύ σημαντική ιδιότητα των νευρωνικών δικτύων πολλών επιπέδων, εν αντιθέσει με τα νευρωνικά δίκτυα απλού Perceptron ενός μόνο επιπέδου. Κάθε δίκτυο MLP έχει την ικανότητα να διαχωρίσει όποια συνεχή επιφάνεια χρειαστεί σε  $n$  διαστάσεις. Κάτι παρόμοιο όμως δεν μπορεί να υλοποιηθεί με τη χρήση ενός απλού δικτύου Perceptron, το οποίο πραγματοποιεί αποκλειστικά και μόνο γραμμικούς διαχωρισμούς.

## 4.2 Αλγόριθμος Error Back Propagation

Όπως και ένα απλό Perceptron ενός επιπέδου, έτσι και το πολυεπίπεδο Perceptron περνάει τη διαδικασία της εκπαίδευσης, στοχεύοντας στην εκμάθηση του δικτύου οποιασδήποτε συνάρτησης είναι απαραίτητο να εκπαιδευτεί και με όποιο βαθμό προσέγγισης. Πρόκειται για τη διαδικασία κατά την οποία πραγματοποιούνται ορισμένες ρυθμίσεις όσον αφορά τις τιμές των βαρών στο νευρωνικό δίκτυο, με σκοπό να ικανοποιηθεί ένα κριτήριο που αφορά την βελτιστοποίηση.



Για την εκπαίδευση ενός δικτύου MLP η πλέον χρησιμοποιούμενη μέθοδος είναι η χρήση του αλγορίθμου πίσω διάδοσης του λάθους. Συγκεκριμένα, ο αλγόριθμος Error Back Propagation στοχεύει στη επίλυση του σφάλματος σύμφωνα με τον κανόνα εκπαίδευσης διόρθωσης του λάθους (error correction learning rule). (Haykin, 2009)

Μια από τις βασικότερες ιδιότητες του συγκεκριμένου αλγορίθμου εκπαίδευσης είναι η έννοια των στόχων. Πιο συγκεκριμένα, αναφερόμαστε στην ύπαρξη στόχων κατά τη διαδικασία μάθησης. Αυτό συνεπάγεται ότι το δίκτυο MLP που χρησιμοποιεί το αλγόριθμο Back Propagation υλοποιεί την επιβλεπόμενη μάθηση.

Είναι ιδιαίτερα σημαντικό να πραγματοποιηθεί μια περιγραφή και ανάλυση της διαδικασίας μάθησης σύμφωνα με τον αλγόριθμο Back Propagation. Για την εκπαίδευση του δικτύου MLP πραγματοποιούνται δύο διαφορετικά περάσματα μέσω των επιπέδων του συγκεκριμένου δικτύου που εκπαιδεύεται.

Σύμφωνα με το πρώτο πέρασμα, το οποίο είναι ένα πέρασμα του δικτύου προς τα εμπρός, πραγματοποιείται μια εισαγωγή ενός διανύσματος στην είσοδο του και συγκεκριμένα στους νευρώνες εισόδου του. Οι επιδράσεις που προκαλούνται στο δίκτυο από την είσοδο του συγκεκριμένου διανύσματος σε αυτό, διαπερνά όλα τα επίπεδα του δικτύου με φορά προς τα εμπρός. Ως τελευταίο στάδιο της διαδικασίας του εμπρός περάσματος είναι η παραγωγή της εξόδου. Η συγκεκριμένη έξοδος που παράγεται αποτελεί την πραγματική απόκριση του MLP δικτύου. Σε όλη τη διαδικασία πραγματοποίησης του πρώτου περάσματος μέσα στο δίκτυο MLP, οι τιμές των βαρών του παραμένουν σταθερές και δε μεταβάλλονται σε κανένα επίπεδο του δικτύου. (Hassan Ramchoun, 2016)

Όσον αφορά το δεύτερο πέρασμα του δικτύου για την εκπαίδευση του πραγματοποιείται ένα πίσω πέρασμα. Σε αυτή τη διαδικασία του περάσματος του δικτύου προς τα πίσω οι τιμές των βαρών του μεταβάλλονται. Η αλλαγή αυτή των τιμών πραγματοποιείται με βάση τον κανόνα « error correction learning rule », που αναφέρθηκε προηγουμένως. Στην περίπτωση αυτή, προκειμένου να υπάρξει ένα σήμα σφάλματος πραγματοποιείται μια αφαίρεση ανάμεσα στην πραγματική με την επιθυμητή έξοδο του δικτύου Perceptron πολλών επιπέδων. Αυτό το σήμα λάθους που έχει δημιουργηθεί διαπερνάει τα επίπεδα του συγκεκριμένου δικτύου με φορά προς τα πίσω. Οι τιμές που έχουν τα συναπτικά βάρη του δικτύου μεταβάλλονται και προσαρμόζονται με σκοπό η πραγματική απόκριση του να φτάσει όσο το δυνατόν περισσότερο την επιθυμητή απόκριση του δικτύου. (Hassan Ramchoun, 2016)

### 4.3 Μειονεκτήματα αλγόριθμου οπισθοδιάδοσης

Ο αλγόριθμος της οπισθοδιάδοσης Back Propagation συνέβαλλε σε πολύ μεγάλο βαθμό στην διαδικασία της επιτυχούς εκπαίδευσης ενός δικτύου Perceptron πολλών επιπέδων. Παρά το γεγονός αυτό όμως ο συγκεκριμένος αλγόριθμος για τη διαδικασία της μάθησης του δικτύου έχει ορισμένα μειονεκτήματα. Πιο συγκεκριμένα, υπάρχουν περιπτώσεις όπου με τη χρήση του αλγόριθμου αυτού η διαδικασία εκπαίδευσης αποτυγχάνεται ή δεν επιφέρει απόλυτη επιτυχία. (Noriega, 2005)

Στις περιπτώσεις αυτές κατά τις οποίες ο αλγόριθμος οπισθοδιάδοσης αποτυγχάνει, το πρόβλημα συνήθως επιλύεται με την αλλαγή των τιμών που έχουν οι παράμετροι του δικτύου MLP καθώς και με την τροποποίηση των αρχικών συνθηκών. Η διαδικασία αυτή της αλλαγής των τιμών του δικτύου συνεχίζεται μέχρις ότου επιλυθεί πλήρως το πρόβλημα της αποτυχίας του αλγόριθμου.

Ένα μειονέκτημα του αλγόριθμου οπισθοδιάδοσης Back Propagation είναι ότι σε ορισμένες περιπτώσεις ο χρόνος που απαιτείται για την ολοκλήρωση της διαδικασίας της μάθησης του δικτύου μπορεί να είναι αρκετά μεγάλος ακόμα και σε υπερβολικό βαθμό. Υπάρχουν περιπτώσεις λόγω χάριν όπου απαιτούνται εκατομμύρια κύκλοι διόρθωσης, προκειμένου η διαδικασία να συγκλίνει. Δεν είναι λίγες οι περιπτώσεις όπου για τόσο μεγάλο αριθμό κύκλων διόρθωσης η διαδικασία δεν φτάνει ποτέ στη σύγκλιση. Σε αυτές τις περιπτώσεις που μόλις αναφέρθηκαν είναι απαραίτητη η τροποποίηση του βήματος μάθησης. Επιπλέον, πρέπει να σημειωθεί ότι σε πολλές περιπτώσεις ακόμα για τεράστιο πλήθος δεδομένων και πληροφοριών, η διαδικασία εκπαίδευσης μπορεί να υλοποιηθεί και να ολοκληρωθεί μέσα σε πολύ μικρό διάστημα μόλις λίγων λεπτών. (Hassan Ramchoun, 2016)

Ένα ακόμα πολύ σημαντικό μειονέκτημα που παρατηρείται στον συγκεκριμένο αλγόριθμο είναι ότι σε μερικές περιπτώσεις υπάρχει η πιθανότητα να συγκλίνει σε ένα τοπικό ελάχιστο εν αντιθέσει με το ολικό της συνάρτησης κόστους που είναι και η επιθυμητή σύγκλιση. Σε γενικές γραμμές τα τοπικά ελάχιστα, αποφεύγονται και δεν θεωρούνται αξιόπιστα διότι αντιμετωπίζονται ως λύσεις χαμηλότερης ακρίβειας σε σχέση με τη λύση που επιφέρει το ολικό ελάχιστο.

Σύμφωνα με το μειονέκτημα αυτό, υπάρχει η περίπτωση η διαδικασία της εκπαίδευσης του δικτύου πολλών επιπέδων MLP να τερματίσει πριν την επιτυχή ολοκλήρωσή της. Με τον τερματισμό της διαδικασίας αυτής, το δίκτυο δεν θα λάβει ως έξοδο το επιθυμητό της αποτέλεσμα.

Προκειμένου να αντιμετωπιστεί σωστά το πρόβλημα αυτό, είναι απαραίτητο όταν ξεκινάει η διαδικασία της εκπαίδευσης του δικτύου οι τιμές των βαρών να αρχικοποιούνται σε πολλές μικρές τιμές. Αυτό συμβαίνει διότι η χρήση μεγάλων τιμών στα συναπτικά βάρη του δικτύου καταλήγουν σε κορεσμό. Όταν επέλθει κορεσμός τότε οι τροποποιήσεις που πραγματοποιούνται στα βάρη είναι πολύ μικρές. Αυτό δεν είναι επιθυμητό διότι το δίκτυο δεν θα φτάσει στο επίπεδο που χρειάζεται για να ολοκληρωθεί επιτυχώς η διαδικασία της μάθησης.

#### 4.4 Τρόποι εκπαίδευσης δικτύου MLP

Όπως έχει ήδη αναφερθεί, ένα δίκτυο Perceptron πολλών επιπέδων MLP εκπαιδεύεται με τη χρήση του αλγόριθμου Πίσω Διάδοσης – Back Propagation. Σύμφωνα με αυτόν τον αλγόριθμο μάθησης γίνεται υλοποίηση στην είσοδο του δικτύου ενός πλήθους από διανύσματα μάθησης. (Haykin, 2009) Με τον όρο Κύκλο ή εποχή μάθησης (epoch) αναφερόμαστε στην ολοκληρωμένη παρουσίαση των διανυσμάτων μάθησης μέσα στο δίκτυο MLP.

Η μάθησης ενός δικτύου MLP εξελίσσεται συνεχώς από κύκλο σε κύκλο. Η διαδικασία της εκπαίδευσης τερματίζει όταν φτάσει η στιγμή που τα συναπτικά βάρη που υπάρχουν στο συγκεκριμένο δίκτυο σταματήσουν να μεταβάλλουν τις τιμές τους και σταθεροποιηθούν και το μέσο τετραγωνικό σφάλμα των διανυσμάτων που χρησιμοποιούνται για τη μάθηση του δικτύου να τείνει σε ένα ελάχιστο.

Διακρίνονται δύο μέθοδοι για την προβολή των διανυσμάτων μάθησης μέσα στο δίκτυο MLP. Οι τρόποι αυτοί είναι οι εξής:

- Στην πρώτη περίπτωση υπάρχει ένας σειριακός τρόπος παρουσίασης των διανυσμάτων μάθησης. Πιο συγκεκριμένα, τα διανύσματα αυτά παρουσιάζονται στο δίκτυο το ένα μετά το άλλο. Η διαδικασία επαναλαμβάνεται μέχρις ότου πραγματοποιηθεί η προβολή όλων των διανυσμάτων ενός κύκλου. Η συγκεκριμένη διαδικασία επαναλαμβάνεται επίσης για όλους τους κύκλους που απαιτούνται.
- Στη δεύτερη περίπτωση η διαδικασία παρουσίασης των διανυσμάτων για την εκπαίδευση του δικτύου πραγματοποιείται με τυχαία σειρά. Με την προϋπόθεση ότι σε κάθε κύκλο που ολοκληρώνεται έχουν παρουσιαστεί οπωσδήποτε, από μία φορά όλα τα δείγματα μάθησης που απαιτούνται. Όπως είναι φανερό, κάθε εποχή εκπαίδευσης διαθέτει κατά πάσα

πιθανότητα διαφορετική σειρά από τους υπόλοιπους κύκλους του δικτύου, δεδομένου ότι πραγματοποιείται τυχαία σειρά παρουσίασης.

Για τη μάθηση ενός δικτύου MLP διακρίνονται δύο τρόποι εκπαίδευσης. Οι τρόποι αυτοί είναι οι εξής:

- **Τρόπος Προτύπων**

Σύμφωνα με τον Τρόπο Προτύπων – Pattern Mode της επιβλεπόμενης μάθησης οι τροποποιήσεις που απαιτούνται στα συναπτικά βάρη του δικτύου πραγματοποιούνται ένα προς ένα. Ο τύπος που υπολογίζει τη συνάρτηση κόστους η οποία θα ελαχιστοποιηθεί για τη συγκεκριμένη περίπτωση είναι ο ακόλουθος: (Hassan Ramchoun, 2016)

$$J = \frac{1}{2} \sum_{i=1}^m \left[ d_i^{(p)} - y_i^{(p)} \right]^2$$

Υπάρχουν N διαφορετικά παραδείγματα για την εκπαίδευση του δικτύου MLP σε κάθε έναν από τους κύκλους μάθησης. Τα διανύσματα αυτά της εκπαίδευσης είναι οργανωμένα με την ακόλουθη σειρά:

$$\{x(1), d(1)\}, \{x(2), d(2)\}, \dots, \{x(N), d(N)\}$$

Σε ένα δίκτυο MLP πολλών επιπέδων το διάνυσμα εκπαίδευσης που εμφανίζετε πρώτο στο δίκτυο είναι το ζεύγος  $\{x(1), d(1)\}$ . Στη συνέχεια, πραγματοποιείται μια τροποποίηση των συναπτικών βαρών του συγκεκριμένου δικτύου. Έπειτα προβάλλεται το δεύτερο κατά σειρά διάνυσμα εκπαίδευσης που είναι το  $\{x(2), d(2)\}$ . Με την προβολή και του δεύτερου παραδείγματος πραγματοποιούνται εκ νέου τροποποιήσεις στα βάρη του δικτύου MLP. Η

διαδικασία που περιγράφηκε θα τερματιστεί όταν παρουσιαστεί στο δίκτυο και το τελευταίο διάνυσμα εκπαίδευσης που είναι το  $\{x(N), d(N)\}$ . Με την παρουσίαση και του τελευταίου ολοκληρώνεται ένας κύκλος εκπαίδευσης. (Hassan Ramchoun, 2016)

- **Μαζική Μάθηση**

Όταν χρησιμοποιείται η μέθοδος της μαζικής μάθησης οι τροποποιήσεις στα συναπτικά βάρη πραγματοποιούνται μετά την ολοκλήρωση της προβολής όλων των διανυσμάτων που απαιτούνται για την εκπαίδευση του δικτύου. Στη μέθοδο αυτή οι τροποποιήσεις πραγματοποιούνται ανά κύκλο μάθησης. Η συνάρτηση κόστους για τον συγκεκριμένο τρόπο εκπαίδευσης η οποία ελαχιστοποιείται είναι η εξής:

$$J = \frac{1}{2N} \sum_{n=1}^N \sum_{i=1}^m \left[ d_i^{(n)} - y_i^{(n)} \right]^2$$

Η μέθοδος εκπαίδευσης του τρόπου προτύπων – Pattern Mode θεωρείται καλύτερη και αποδοτικότερη από τη μαζική μάθηση – Batch Mode. Ο λόγος είναι διότι ο τρόπος των προτύπων απαιτεί μικρότερο χώρο για την αποθήκευση των συνάψεων του δικτύου. Ένας ακόμα πολύ σημαντικός λόγος είναι ότι είναι πολύ πιο εύκολο να παρακολουθούνται οι μικρές τροποποιήσεις στις πληροφορίες της εκπαίδευσης εν αντιθέσει με τις μεγάλες και πολλές τροποποιήσεις. (Hassan Ramchoun, 2016)

# ΚΕΦΑΛΑΙΟ 5 – ΑΛΓΟΡΙΘΜΟΣ ΜΕΣΟΥ ΤΕΤΡΑΓΩΝΙΚΟΥ ΣΦΑΛΜΑΤΟΣ

---

## 5.1 Εισαγωγή

Το συγκεκριμένο κεφάλαιο ασχολείται με την ανάλυση του αλγορίθμου του μέσου τετραγωνικού σφάλματος (Least Mean Square Algorithm - LMS). Ο αλγόριθμος αυτός δημιουργήθηκε το 1960 από τους Hoff και Widrow. (Haykin, 2009)

Το μοντέλο του Perceptron, το οποίο δημιουργήθηκε από τον Rosenblatt έχει περιγραφεί αναλυτικότερα στο κεφάλαιο 3 της συγκεκριμένης εργασίας. Όπως έχει ήδη αναφερθεί πρόκειται για το πιο απλό μοντέλο νευρωνικού δικτύου, το οποίο αποτελείται από έναν και μόνο νευρώνα. Υπήρξε ο πρώτος αλγόριθμος για τη μάθηση, ο οποίος στόχευε στην επίλυση γραμμικά διαχωρίσιμων προβλημάτων για την ταξινόμηση προτύπων.

Από την άλλη μεριά ο αλγόριθμος μέσου τετραγωνικού σφάλματος – Least Mean Square Algorithm (LMS) αποτελεί τον πρώτο γραμμικό αλγόριθμο που στοχεύει στην προσαρμογή καθώς και το φιλτράρισμα προκειμένου να επιλύει διάφορα είδη προβλημάτων. Η ιδέα της δημιουργίας και της υλοποίησης του αλγορίθμου αυτού βασίστηκε σε πολύ μεγάλο βαθμό στο δίκτυο Perceptron, που έχει ήδη περιγραφεί. (Haykin, 2009)

Αν και διακρίνονται ορισμένες διαφορές ανάμεσα στον αλγόριθμο του Perceptron και στον αλγόριθμο του μέσου τετραγωνικού σφάλματος όσον αφορά την εφαρμογή τους, υπάρχει ένα βασικό χαρακτηριστικό που διαθέτουν και οι δύο αλγόριθμοι. Το κοινό αυτό χαρακτηριστικό σχετίζεται με τη χρήση ενός γραμμικού συνδυαστή (linear combiner). (Haykin, 2009)

Αξιοσημείωτο είναι το γεγονός ότι ο αλγόριθμος LMS που μελετάται στο κεφάλαιο αυτό δε χρησιμοποιείται αποκλειστικά και μόνο για τη χρήση του σε εφαρμογές φιλτραρίσματος και προσαρμογής αλλά συνεισφέρει σημαντικά ως σημείο αναφοράς όσον αφορά την αξιολόγηση άλλων αλγορίθμων προσαρμογής – φιλτραρίσματος. (Haykin, 2009)

Στην συνέχεια αναφέρονται οι λόγοι που οδήγησαν στην εξέλιξη του αλγορίθμου του μέσου τετραγωνικού σφάλματος LMS καθώς και το γεγονός ότι μπορεί να χρησιμοποιηθεί και για την αξιολόγηση άλλων αλγορίθμων προσαρμογής – φιλτραρίσματος.

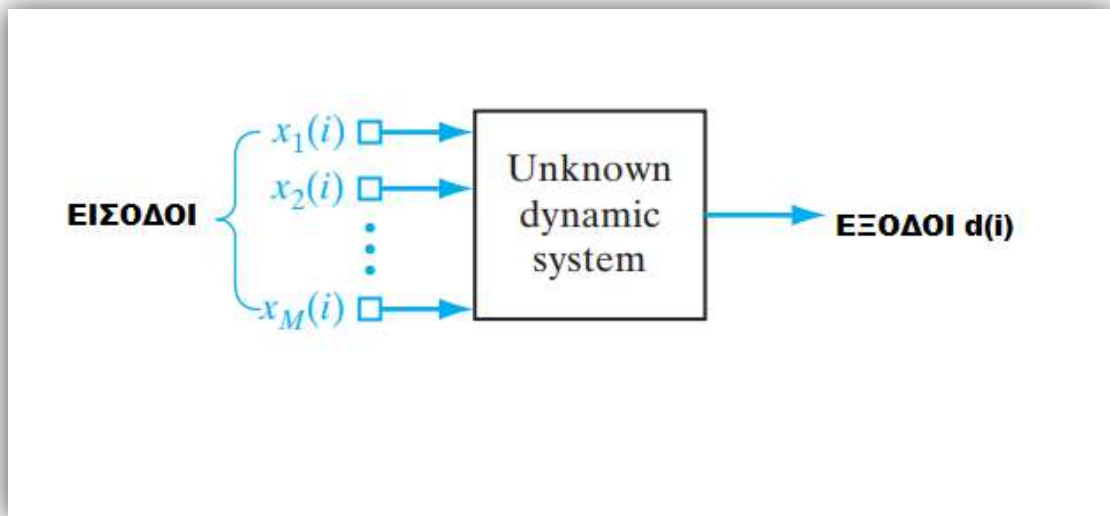
- Ο αλγόριθμος μέσου τετραγωνικού σφάλματος θεωρείται ένας απλός αλγόριθμος όσον αφορά τον κώδικά του. Βασιζόμενοι σε αυτό το γεγονός, καταλήγουμε στο συμπέρασμα ότι ο αλγόριθμος LMS είναι ένας εύκολος αλγόριθμος ως προς την κατασκευή του.
- Όσον αφορά την πολυπλοκότητα του αλγορίθμου, αυτή είναι γραμμική σε σχέση με τις παραμέτρους του.
- Ένα πολύ σημαντικό χαρακτηριστικό του αλγορίθμου αυτού που το καθιστά ένα πολύτιμο εργαλείο είναι η ισχύς του. Πρόκειται για έναν πολύ ισχυρό αλγόριθμο όσον αφορά τις εξωτερικές μεταβολές και διαταραχές που μπορούν να προκύψουν. (Haykin, 2009)

Όλες οι ιδιότητες που περιγράφηκαν προηγουμένως και που διαθέτει ο αλγόριθμος μέσου τετραγωνικού σφάλματος LMS, θεωρούνται πολύ σημαντικές ιδιότητες. Η ύπαρξη τους σε όλους τους αλγορίθμους είναι ιδιαίτερα επιθυμητή. Αυτός είναι και ο λόγος για τον οποίο ο αλγόριθμος LMS έχει παραμείνει στο προσκήνιο με την πάροδο του χρόνου. (Haykin, 2009)

## 5.2 Διαδικασία Φιλτραρίσματος του LMS αλγορίθμου

Στο Σχήμα 19 που βρίσκεται παρακάτω απεικονίζεται ένα διάγραμμα, αναπαριστώντας τη δομή ενός αγνώστου δυναμικού συστήματος. Το σύστημα αυτό δέχεται ως είσοδο ένα διάνυσμα, το οποίο απαρτίζεται από τα στοιχεία  $x_1(i)$ ,  $x_2(i)$ , ...,  $x_M(i)$ , όπως φαίνεται και από την παρακάτω εικόνα. Με το σύμβολο  $i$  συμβολίζεται η χρονική στιγμή στην οποία το σύστημα διεγείρεται. Μετά από την διέγερση του συστήματος, αυτό με τη σειρά του παράγει μια έξοδο, η οποία συμβολίζεται ως  $y(i)$ . (Haykin, 2009)

Ένα πρότυπο του δυναμικού συστήματος, το οποίο αποτελείται από πολλές εισόδους και παράγει μια μόνο έξοδο καλείτε να αντιμετωπίσει και να επιλύσει προβλήματα με τη χρήση ενός μόνο νευρώνα. Με τη χρήση του αλγορίθμου, πραγματοποιούνται αλλαγές και τροποποιήσεις στις τιμές των συναπτικών βαρών του μοναδικού νευρώνα που υπάρχει στο σύστημα.



Εικόνα 26 - Δυναμικό Σύστημα

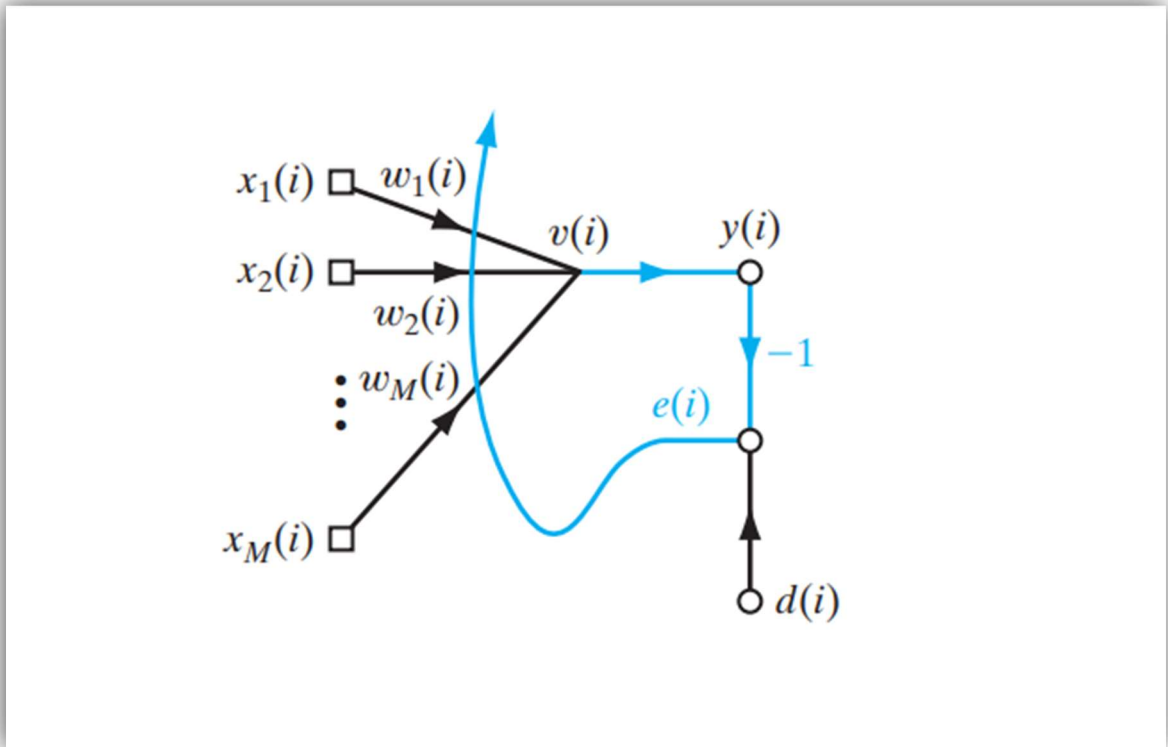
Οι αλλαγές που πραγματοποιούνται στα συναπτικά βάρη του νευρώνα που υπάρχει στο σύστημα βασίζονται στα εξής:

- Αρχικά οι αλλαγές που πραγματοποιούνται από τον αλγόριθμο όσον αφορά τα συναπτικά βάρη γίνονται αυθαίρετα.
- Οι προσαρμογές στις τιμές των συναπτικών βαρών μέσα στο άγνωστο δυναμικό σύστημα πραγματοποιούνται σε συνεχή βάση.
- Όλοι οι υπολογισμοί που υλοποιούνται στα συναπτικά βάρη ολοκληρώνονται μέσα σε μια περίοδο δειγματοληψίας.

Η διαδικασία που μόλις περιγράφηκε αναφέρεται σε ένα προσαρμοστικό φίλτρο (adaptive filter). (Haykin, 2009)



Στην Εικόνα 20 που ακολουθεί απεικονίζετε ένας γράφος, ο οποίος αναφέρεται στη ροή του σήματος ενός προσαρμοστικού φίλτρου. Με μπλε χρώμα αναπαριστάνεται έναν βρόχο ανατροφοδότησης που υπάρχει στο συγκεκριμένο σύστημα.



Εικόνα 27 - Γράφημα ροής σήματος προσαρμοστικού μοντέλου

Η λειτουργία ενός προσαρμοστικού φίλτρου βασίζεται σε δύο διαφορετικές διαδικασίες οι οποίες είναι οι εξής: (Haykin, 2009)

- **Διαδικασία Φιλτραρίσματος**

Η πρώτη διαδικασία, η οποία είναι η διαδικασία φιλτραρίσματος αποτελείται από δύο διαφορετικούς υπολογισμούς σημάτων. Το πρώτο σήμα που καλείται να υπολογίσει είναι το σήμα εξόδου. Όπως έχει ήδη αναφερθεί η έξοδος του συστήματος συμβολίζεται ως  $y(i)$ .

Το δεύτερο σήμα που πρέπει να υπολογιστεί είναι το σήμα σφάλματος και συμβολίζεται ως  $e(i)$ . Το συγκεκριμένο σήμα υπολογίζεται συγκρίνοντας την έξοδο  $y(i)$  με την έξοδο που παράγεται από το άγνωστο δυναμικό σύστημα που συμβολίζεται με το  $d(i)$ . Η έξοδος του άγνωστου συστήματος θεωρείται ως το επιθυμητό σήμα απόκρισης.

- **Διαδικασία Προσαρμογής**

Η δεύτερη διαδικασία είναι η προσαρμοστική. Πρόκειται για μια αυτόματη προσαρμογή στις τιμές των βαρών του νευρώνα που υπάρχει στο κάθε σύστημα. Η συγκεκριμένη προσαρμογή βασίζεται στην τιμή που έχει το σήμα του σφάλματος και συμβολίζεται ως  $e(i)$ .

Οι δύο αυτές διαδικασίες που περιγράφηκαν προηγουμένως και συγκεκριμένα η διαδικασία του Φιλτραρίσματος και η διαδικασία της Προσαρμογής συνδυάζονται και λειτουργούν αρμονικά μεταξύ τους ώστε να πραγματοποιηθεί ένας βρόγχος που λειτουργεί γύρω από τον νευρώνα που υπάρχει στο δίκτυο. Στην εικόνα 20 αναπαριστάνεται και ο συγκεκριμένος βρόγχος.

Έχοντας ως δεδομένο ότι ο νευρώνας που ασχολούμαστε στο συγκεκριμένο κεφάλαιο είναι ένας γραμμικός νευρώνας, παρατηρούμε ότι η έξοδος του είναι η ίδια με την τιμή του  $u(i)$

Συγκεκριμένα η τιμή της εξόδου του νευρώνα υπολογίζεται από τον ακόλουθο τύπο:

$$y(i) = u(i) = \sum_{k=1}^M w_k(i) x_k(i)$$

Στον παραπάνω τύπο τα  $w_1(i)$ ,  $w_2(i)$ , ...,  $w_M(i)$  συμβολίζουν τα συναπτικά βάρη που υπάρχουν τη χρονική στιγμή  $i$  στον νευρώνα του δικτύου.

Μπορούμε να εκφράσουμε την έξοδο του δικτύου σε μορφή πίνακα, με τη χρήση του εσωτερικού γινομένου των  $x(i)$  καθώς και  $w(i)$ . Το εσωτερικό αυτό γινόμενο βρίσκεται στη συνέχεια:

$$y(i) = x^T(i) w(i)$$

Όπου το  $w(i)$  που βρίσκεται στον παραπάνω τύπο υπολογίζεται με τον ακόλουθο τύπο:

$$w(i) = [w_1(i), w_2(i), \dots, w_M(i)]^T$$

Με τον όρο σφάλμα  $e$  του συστήματος αναφερόμαστε στη διαφορά που υπάρχει ανάμεσα στην πραγματική έξοδο που υπολογίζεται από το σύστημα  $y(i)$  και στην επιθυμητή έξοδο  $d(i)$ . Πιο συγκεκριμένα, από τη σύγκριση της εξόδου του νευρώνα και την αντίστοιχη έξοδο  $d(i)$  που υπολογίζεται τη χρονική στιγμή  $i$  μέσα από ένα άγνωστο σύστημα λαμβάνεται το σφάλμα. Ο τύπος που υπολογίζει το σήμα σφάλματος είναι ο εξής:

$$e(i) = d(i) - y(i)$$

### 5.3 Βελτιστοποίηση χωρίς περιορισμούς

Το σήμα σφάλματος  $e$  (i) χρησιμοποιείται με τέτοιο τρόπο ώστε να πραγματοποιείται ο έλεγχος των μεταβολών στα συναπτικά βάρη του νευρώνα που καθορίζεται από τη συνάρτηση κόστους. Η διαδικασία αυτή είναι άμεσα συνδεδεμένη με τη βελτιστοποίηση. Είναι εξίσου σημαντικό να πραγματοποιηθεί μια αναφορά στην βελτιστοποίηση χωρίς περιορισμούς.

Αρχικά εξετάζετε μια συνάρτηση κόστους την οποία συμβολίζουμε ως  $e(w)$ . Σκοπός είναι να βρεθεί μια βέλτιστη λύση  $w^*$ . Η λύση αυτή πρέπει να ικανοποιεί τον ακόλουθο κανόνα:

$$e(w^*) \leq e(w)$$

Καλούμαστε να λύσουμε ένα πρόβλημα βελτιστοποίησης. Υπάρχει μια απαραίτητη προϋπόθεση όσον αφορά την βελτιστοποίηση. Η προϋπόθεση αυτή είναι η ακόλουθη:

$$\nabla e(w^*) = 0$$

Η διαδικασία ξεκινάει από μια αρχική υπόθεση όσον αφορά το  $w(0)$ , με σκοπό να δημιουργηθεί η ακολουθία των βαρών  $w(1)$ ,  $w(2)$ . Η συγκεκριμένη διαδικασία έχει ως στόχο την μείωση της συνάρτησης κόστους σε κάθε μια από τις επαναλήψεις του αλγορίθμου. Αυτό επιτυγχάνεται όταν:

$$\mathcal{E}(w(n+1)) < \mathcal{E}(w(n))$$

Διακρίνονται τρεις διαφορετικές μέθοδοι βελτιστοποίησης χωρίς περιορισμούς. Πιο συγκεκριμένα, οι μέθοδοι αυτοί είναι η μέθοδος απότομης καθόδου (Method of Steepest Descent), η μέθοδος Newton καθώς και η μέθοδος Gauss-Newton. (Haykin, 2009)

#### 5.4 Αλγόριθμος Least Mean Square - LMS

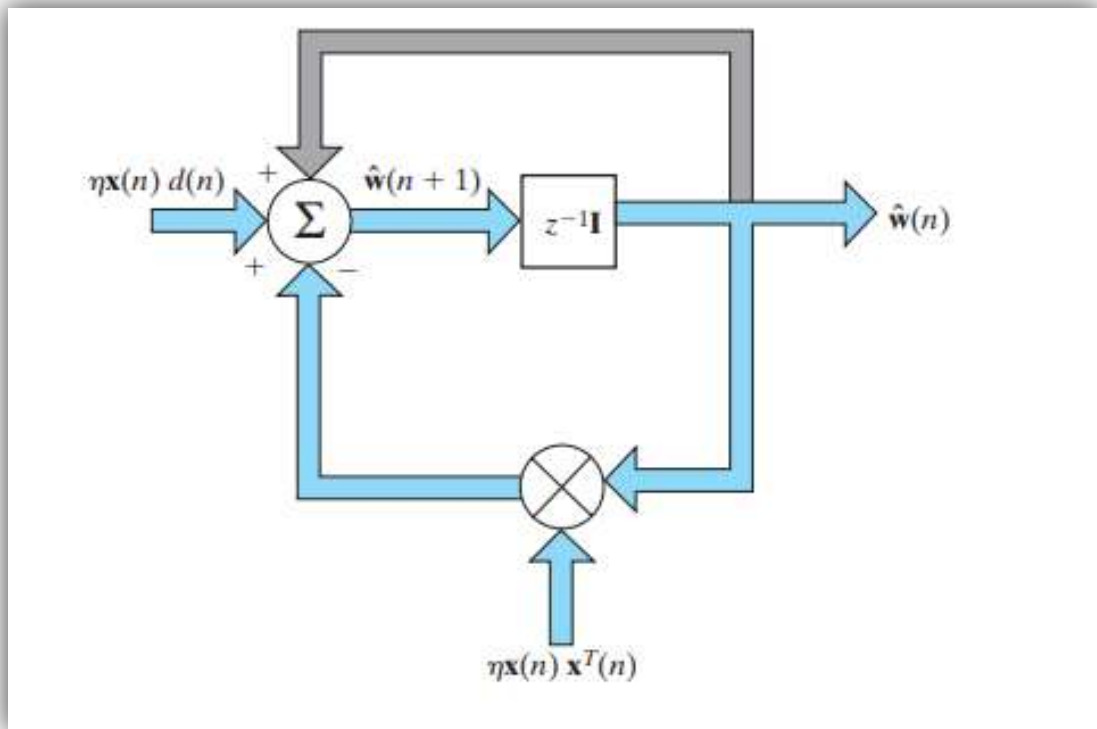
Ο κύριος στόχος της χρήσης του αλγορίθμου του μέσου τετραγωνικού σφάλματος – Least mean square είναι η ελαχιστοποίηση της τιμής που έχει η συνάρτηση κόστους.

$$\mathcal{E}(\hat{w}) = \frac{1}{2} e^2(n)$$

Όπως έχει ήδη αναφερθεί το  $e$  συμβολίζει το σήμα σφάλματος σε χρονικό διάστημα  $n$ . Ο συγκεκριμένος αλγόριθμος που μελετάται λειτουργεί με έναν γραμμικό νευρώνα. Για το λόγο αυτό, το σήμα σφάλματος μπορεί να υπολογιστεί και από τον ακόλουθο τύπο:

$$e(n) = d(n) - x^T(n) \hat{w}(n)$$

Στην εικόνα 22 που βρίσκεται στη συνέχεια απεικονίζεται η ροή του σήματος όπως αυτό συμβαίνει στον αλγόριθμο Μέσου τετραγωνικού σφάλματος LMS.



Εικόνα 28 - Ροή Σήματος του αλγορίθμου LMS

Στη συνέχεια περιγράφονται τα βήματα της διαδικασίας που ακολουθεί ο αλγόριθμος του μέσου τετραγωνικού σφάλματος για την υλοποίησή του:

- Το πρώτο βήμα αποτελεί η αρχικοποίηση. Αρχικά θέτει το  $\hat{w}_k(1)$  ίσο με τη μηδενική τιμή.

$$\hat{w}_k(1) = 0 \text{ για } k = 1, 2, \dots, p$$

- Στο δεύτερο βήμα του αλγορίθμου πραγματοποιούνται οι ακόλουθοι υπολογισμοί:

$$y(n) = \sum_{j=1}^p \hat{w}_j(n) x_j(n)$$

και

$$e(n) = d(n) - y(n)$$

Και τέλος εκτελείται ο ακόλουθος υπολογισμός:

$$\hat{w}_k(n+1) = w_k n + \eta e(n) x_k(n) \text{ για } k= 1, 2, \dots, p$$

Όπως αναφέραμε παραπάνω πραγματοποιείται αρχικά μια αρχικοποίηση του αλγορίθμου. Η τιμή που παίρνει η αρχικοποίηση των βαρών συνηθίζεται να είναι η μηδενική τιμή.

# ΚΕΦΑΛΑΙΟ 6 – ΣΥΜΠΕΡΑΣΜΑΤΑ

---

## 6.1 Συμπεράσματα

Στα πλαίσια της συγκεκριμένης πτυχιακής εργασία πραγματοποιείτε μια περιγραφή και ανάλυση του κλάδου των Τεχνητών Νευρωνικών Δικτύων. Μελετήθηκε η λειτουργία των Τεχνητών Νευρωνικών Δικτύων καθώς και η λειτουργία των Βιολογικών Νευρωνικών Δικτύων που είναι ο ανθρώπινος εγκέφαλος. Πραγματοποιήθηκε μια ιστορική αναφορά καθώς και μια περιγραφή των εφαρμογών που υπάρχουν και χρησιμοποιούν τα Τεχνητά Νευρωνικά Δίκτυα.

Ο λόγος για τον οποίο οι επιστήμονες του κλάδου των Νευρωνικών δικτύων επιδίωξαν να δημιουργήσουν τα Τεχνητά νευρωνικά δίκτυα ήταν για να μπορέσουν να προσομοιώσουν όσο το δυνατόν καλύτερα τις πολύπλοκες λειτουργίες που επιτελεί ο ανθρώπινος εγκέφαλος.

Τα Τεχνητά Νευρωνικά Δίκτυα αποτελούν έναν καινούργιο επιστημονικό κλάδο, ο οποίος αναπτύχθηκε περαιτέρω τα τελευταία χρόνια. Πρόκειται για έναν πολλά υποσχόμενο κλάδο των επιστημών, ικανός να επιλύσει προβλήματα από πολλούς και διαφορετικούς επιστημονικούς κλάδους, όπως λόγου χάριν την ιατρική, την οικονομία και την επιστήμη των υπολογιστών.

Επίσης, στην εργασία αυτή αναφέρθηκαν και αναλύθηκαν περαιτέρω οι δύο διαφορετικές Αρχιτεκτονικές των νευρωνικών δικτύων που υπάρχουν. Συγκεκριμένα, περιγράφηκαν τα Νευρωνικά Δίκτυα πρόσθιας τροφοδότησης (feedforward) καθώς και τα Νευρωνικά Δίκτυα με ανατροφοδότηση (feedback, recurrent).

Πιο αναλυτικά, στα Τεχνητά Νευρωνικά δίκτυα που χρησιμοποιούν την πρόσθια τροφοδότηση, τα σήματα διαδίδονται αποκλειστικά και μόνο με κατεύθυνση από την είσοδο του νευρωνικού δικτύου προς την έξοδο. Ενώ από την άλλη μεριά στα Νευρωνικά Δίκτυα που χρησιμοποιούν ως αρχιτεκτονική την ανατροφοδότηση, τα σήματα διαδίδονται και στις δύο κατευθύνσεις, μέσω των βρόγχων που υπάρχουν στο δίκτυο.

Ύστερα από μια γενική και βασική ανάλυση των Δικτύων αυτών, ασχοληθήκαμε με τη Μηχανική Μάθηση. Ο λόγος που προέτρεψε τους επιστήμονες να κατασκευάσουν τα τεχνητά νευρωνικά δίκτυα ήταν η δημιουργία συστημάτων, τα οποία θα μπορούσαν να επιλύσουν διάφορες διεργασίες καθώς και να εκτελούν από μόνα τους μια λειτουργία. Προκειμένου να συμβεί κάτι ανάλογο θα πρέπει το δίκτυο να εκπαιδευτεί σωστά.



Όπως έχει ήδη αναφερθεί, υπάρχουν τρία είδη μάθησης των νευρωνικών δικτύων. Πιο συγκεκριμένα, διακρίνουμε την Επιβλεπόμενη Μάθηση (Supervised Learning), την Ενισχυτική Μάθηση (Reinforcement Learning) και την Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning).

Μπορούμε να αναφέρουμε τη μάθηση με επίβλεψη και ως μάθηση με εκπαιδευτή. Αυτός ο τρόπος μάθησης αποτελεί τον πιο συχνά χρησιμοποιούμενο τρόπο για την εκπαίδευση των Τεχνητών νευρωνικών δικτύων. Διακρίνονται δύο διαφορετικοί τύποι της ενισχυτικής μάθησης. Ο πρώτος τύπος είναι η συσχετική ενισχυτική μάθηση. Σύμφωνα με το τύπο αυτό, εκτός από το ενισχυτικό σήμα που τροφοδοτείται στο Νευρωνικό Δίκτυο, τροφοδοτούνται επίσης διάφοροι τύποι πληροφοριών. Ο δεύτερος τύπος είναι η μη συσχετιστική ενισχυτική μάθηση. Στον τύπο αυτό, η μόνη είσοδος από το περιβάλλον στο νευρωνικό δίκτυο είναι το ενισχυτικό σήμα.

Στην προσπάθεια για πλήρη κατανόηση του κλάδου των Τεχνητών Νευρωνικών Δικτύων πραγματοποιήθηκε μια ανάλυση τόσο του απλού δικτύου Perceptron που αποτελείται από έναν και μόνο νευρώνα, όσο και του πολυεπίπεδου δικτύου Perceptron που αποτελείται από περισσότερους νευρώνες.

Το απλό δίκτυο Perceptron ενός επιπέδου αποτελεί το πρώτο μοντέλο Τεχνητού Νευρωνικού Δικτύου το οποίο δημιουργήθηκε και αποτέλεσε βάση για την μετέπειτα ανάπτυξη πιο περίπλοκων και σύνθετων δικτύων. Ο κύριος στόχος για τη δημιουργία του δικτύου Perceptron είναι να χρησιμοποιείται κατάλληλα προκειμένου να επιτυγχάνεται η ταξινόμηση των προτύπων. Προκειμένου να πετύχει τη σωστή λειτουργία του, θα πρέπει να χρησιμοποιείται ο λεγόμενος γραμμικός διαχωρισμός. Για τη διαδικασία μάθησης του Νευρωνικού Δικτύου Perceptron χρησιμοποιείτε η εκπαίδευση με επίβλεψη.

Υπάρχει ένα πολύ σημαντικό μειονέκτημα που παρατηρείται σε αυτό το δίκτυο Perceptron. Αναφερόμαστε στο γεγονός ότι δε μπορούν με το συγκεκριμένο δίκτυο να επιλυθούν μη γραμμικά διαχωρίσιμα προβλήματα. Όλα σχεδόν τα προβλήματα που υπάρχουν στην πραγματικότητα είναι μη γραμμικά διαχωρίσιμα. Για το λόγο αυτό, οι επιστήμονες που ασχολήθηκαν με τα Τεχνητά Νευρωνικά Δίκτυα πραγματοποίησαν μια προσπάθεια εξέλιξης του απλού Perceptron, στη δημιουργία του πολυεπίπεδου Perceptron (Multi-Layer Perceptron - MLP). Πρόκειται ουσιαστικά για μια γενίκευση του απλού δικτύου Perceptron.

Πιο συγκεκριμένα, ένα δίκτυο Perceptron πολλών διαθέτει ένα πλήθος αισθητήρων που υπάρχουν στο επίπεδο της εισόδου καθώς και ένα πλήθος υπολογιστικών κόμβων που αποτελούν το επίπεδο της εξόδου. Ανάμεσα από τα δύο αυτά επίπεδα υπάρχουν ένα ή και περισσότερα κρυφά επίπεδα (hidden layers).

Για την εκπαίδευση ενός δικτύου MLP η πλέον χρησιμοποιούμενη μέθοδος είναι η χρήση του αλγορίθμου πίσω διάδοσης του λάθους. Συγκεκριμένα, ο αλγόριθμος Error Back Propagation στοχεύει στη επίλυση του σφάλματος σύμφωνα με τον κανόνα εκπαίδευσης διόρθωσης του λάθους.

Ως τελευταίο αντικείμενο μελέτης της συγκεκριμένης πτυχιακής εργασίας είναι η ανάλυση του αλγορίθμου του μέσου τετραγωνικού σφάλματος (Least Mean Square Algorithm - LMS). Ο αλγόριθμος αυτός δημιουργήθηκε το 1960 από τους Hoff και Widrow.

## 6.2 Μελλοντικές εξελίξεις των Νευρωνικών Δικτύων

Ο κλάδος των Τεχνητών Νευρωνικών Δικτύων έχει δημιουργηθεί και αναπτυχθεί περαιτέρω τα τελευταία χρόνια. Πρόκειται για έναν σχετικά καινούργιο κλάδο της επιστήμης. Ο κύριος λόγος που οδήγησε τους επιστήμονες στη δημιουργία των Τεχνητών Νευρωνικών Δικτύων είναι η προσπάθεια αναπαράστασης της λειτουργίας του ανθρώπινου εγκεφάλου που αποτελεί ένα Βιολογικό Νευρωνικό Δίκτυο.

Είναι ένας κλάδος που έχει προσφέρει πολλές και χρήσιμες λειτουργίες στους επιστήμονες. Αναμένεται, επίσης να αναπτυχθεί περαιτέρω μέσα στα επόμενα χρόνια που έρχονται και να προσφέρει επιπλέον δυνατότητες στους επιστημονικούς κλάδους.

Αν και η μέχρι τώρα πορεία του κλάδου των Νευρωνικών Δικτύων έχει παρουσιάσει ορισμένα μειονεκτήματα ως προς τις λειτουργίες τους, τα πλεονεκτήματα τους είναι πολλά περισσότερα. Οι επιστήμονες του συγκεκριμένου κλάδου έχουν διευκρινίσει την δυνατότητα της μελλοντικής εξέλιξης και βελτίωσης των λειτουργιών των Νευρωνικών Δικτύων.

Από τη συνεχή έρευνα και μελέτη των Τεχνητών Νευρωνικών Δικτύων από τους επιστήμονες του κλάδου πραγματοποιούνται συνεχώς προσπάθειες βελτίωσης και εξέλιξης τους. Μελετώντας διεξοδικά τις πολύπλοκες λειτουργίες του ανθρώπινου εγκεφάλου προσπαθούν να τις προσομοιώσουν στα Τεχνητά Νευρωνικά Δίκτυα.

Οι επιστήμονες στοχεύουν στην κατασκευή πιο εξελιγμένων δικτύων με τη χρήση πιο ρεαλιστικών αλγορίθμων για τη μάθηση των δικτύων αυτών καθώς και πιο ισχυρών τρόπων όσον αφορά την σύνδεση των νευρώνων που υπάρχουν στο νευρωνικών δίκτυο.

Οι λειτουργίες των Νευρωνικών δικτύων πρέπει στο μέλλον να εξελιχθούν προκειμένου να χρησιμοποιηθούν για πιο πολύπλοκες και απαιτητικές εφαρμογές. Αυτό μπορεί να επιτευχθεί με την συνεχή βελτίωση της ισχύς που διαθέτουν οι ηλεκτρονικοί υπολογιστές όσον αφορά την υπολογιστική τους λειτουργία. Βελτίωση της υπολογιστικής ισχύς σε συνδυασμό με τη δημιουργία Τεχνητών Νευρωνικών Δικτύων σε ολοκληρωμένα κυκλώματα χρησιμοποιώντας την VLSI τεχνολογία μπορούν να επιτύχουν βελτίωση των δυνατοτήτων του κλάδου των Νευρωνικών Δικτύων.

# ΒΙΒΛΙΟΓΡΑΦΙΑ

---

- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*.
- D. Anderson, G. M. (1992). *Artificial Neural Networks Technology*.
- GALLANT, S. (1990). Perceptron-based learning algorithms.
- Geoffrey E. Hinton, T. J. (1999). *Unsupervised Learning: Foundations of Neural Computation*.
- Hassan Ramchoun, M. A. (2016). *Multilayer Perceptron: Architecture*.
- Haykin, S. (2009). *Neural Networks and Learning Machines*.
- Jantzen, J. (1998). Retrieved from Introduction To Perceptron Networks.
- Noriega, L. (2005, November 17). *Multilayer Perceptron Tutorial*.
- P.J. Braspenning, A.J.M.M. Weijters, F. Thuijsman. (1995 ). *Artificial Neural Networks- An Introduction to ANN Theory and Practice*.
- Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*.
- Rojas, R. (1996). *Neural Networks - A systematic Introdustion*. Berlin: Springer-Verlag.
- Russell Reed, R. J. (1999). *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*.
- Shepherd, A. J. (1997). *Second-Order Methods for Neural Networks: Fast and Reliable Training Methods*.
- Sutton, R. (1998). *Introduction to Reinforcement Learning*.
- W. Thomas Miller, R. S. (1990). *Neural Networks for Control*.
- WIDROW, B. (1990). 30 Years of Adaptive Neural Networks : Perceptron, Madaline, and Backpropagation.
- Widrow, B. (1998). *The Handbook of Brain Theory and Neural Networks. Percetrons, Adalines and Backpropagation*.
- Wilde, P. d. (1996). *Neural Networks Models: Theory and Projects*.

Αργυράκης, Π. (2001). *Νευρωνικά Δίκτυα και Εφαρμογές*. ΕΛΛΗΝΙΚΟ ΑΝΟΙΚΤΟ ΠΑΝΕΠΙΣΤΗΜΙΟ.

Ν Ζώρης, Ν. Κ. (2014). *Νευρωνικά δίκτυα και εφαρμογές*.

Πλήρου, Α. (2012). *Artificial neural networks simulating human brain*. Ελληνικό Ανοικτό Πανεπιστήμιο.