

**Σελίδα τίτλου**

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**ΕΚΠΑΙΔΕΥΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΜΕ ΣΤΟΧΑΣΤΙΚΕΣ ΤΕΧΝΙΚΕΣ**

Παρασκευή Τοκμακίδου

Επιβλέπων: Ιωάννης Τσούλος

τίτλος, βαθμίδα

Τόπος έκδοσης, Μήνας, Έτος ολοκλήρωσης

**Ο ΤΙΤΛΟΣ ΤΗΣ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ ΜΕ ΚΕΦΑΛΑΙΑ ΓΡΑΜΜΑΤΑ ΣΤΗΝ ΑΓΓΛΙΚΗ ΓΛΩΣΣΑ ΚΑΙ ΣΤΟΙΧΙΣΗ ΣΤΟ ΚΕΝΤΡΟ**

**Εγκρίθηκε από τριμελή εξεταστική επιτροπή**

Τόπος, Ημερομηνία

**ΕΠΙΤΡΟΠΗ ΑΞΙΟΛΟΓΗΣΗΣ**

1. Επιβλέπων καθηγητής

Όνοµα Επίθετο,

1. Μέλος επιτροπής

Όνοµα Επίθετο,

1. Μέλος επιτροπής

Όνοµα Επίθετο,

**©** Τοκμακίδου, Παρασκευή, 2024.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

**Δήλωση μη λογοκλοπής**

Δηλώνω υπεύθυνα και γνωρίζοντας τις κυρώσεις του Ν. 2121/1993 περί Πνευματικής Ιδιοκτησίας, ότι η παρούσα πτυχιακή εργασία είναι εξ ολοκλήρου αποτέλεσμα δικής μου ερευνητικής εργασίας, δεν αποτελεί προϊόν αντιγραφής ούτε προέρχεται από ανάθεση σε τρίτους. Όλες οι πηγές που χρησιμοποιήθηκαν (κάθε είδους, μορφής και προέλευσης) για τη συγγραφή της περιλαμβάνονται στη βιβλιογραφία.

Τοκμακίδου, Παρασκευή

Υπογραφή

# ΕΥΧΑΡΙΣΤΙΕΣ

Είναι σημαντική η αναγνώριση της βοήθειας που έλαβε ο φοιτητής/ η φοιτήτρια κατά τη διάρκεια της προπαρασκευής της εργασίας του. Η βοήθεια μπορεί να είναι ακαδημαϊκή, τεχνική, γραμματειακή, διοικητική και προσωπική (π.χ. οικογένεια). Δεν υπερβαίνει τη μία παράγραφο.

*[Διαγράψτε αυτή τη σελίδα αν δεν τη χρειάζεστε.]*

Στην συγγραφή αυτής της πτυχιακής εργασίας είχα την πρακτική και ηθική υποστήριξη διάφορων συνεργατών και φίλων. Τέλος, θέλω να ευχαριστήσω την οικογένεια μου που έκαναν υπομονή καθώς αφιέρωνα σημαντικό κομμάτι του χρόνου μου για την ολοκλήρωση του έργου αυτού.

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή κ. Τσούλο Ιωάννη για την δυνατότητα

που μου έδωσε να πραγματοποιήσω την παρούσα εργασία όπως σε ένα τέτοιο ενδιαφέρον

τομέα όπως είναι τα Νευρωνικά Δίκτυα καθώς και για την σημαντική βοήθειά του.

Αφιερώνεται στον/στην

για την υπομονή και την ανεκτικότητα της, καθώς και στους αμέτρητους ερευνητές που ασχολούνται με το πεδίο των νευρωνικών δικτύων.

# ΠΕΡΙΛΗΨΗ

Η περίληψη (στην ελληνική γλώσσα) αποτελεί μια συνοπτική παρουσίαση των κύριων στοιχείων και συμπερασμάτων της εργασίας, μαζί με μια σύντομη αναφορά στους στόχους και τις μεθοδολογίες που ακολουθήθηκαν. Δεν θα πρέπει να υπερβαίνει τις 300 λέξεις.

**Λέξεις-κλειδιά**: Καταχωρίστε τρεις έως πέντε λέξεις ή φράσεις-κλειδιά στα ελληνικά που περιγράφουν το θέμα της εργασίας.

# ABSTRACT

Η περίληψη στην αγγλική γλώσσα. Δεν θα πρέπει να υπερβαίνει τις 300 λέξεις.

**Keywords**: Καταχωρίστε τρεις έως πέντε λέξεις ή φράσεις-κλειδιά στα αγγλικά που περιγράφουν το θέμα της εργασίας.

# ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

[ΕΥΧΑΡΙΣΤΙΕΣ 9](#_Toc30481)

[ΠΕΡΙΛΗΨΗ 10](#_Toc25670)

[ABSTRACT 11](#_Toc16026)

[ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ 12](#_Toc20655)

[ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ 14](#_Toc25198)

[ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ 15](#_Toc5035)

[ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ 16](#_Toc20790)

[ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ 17](#_Toc22203)

[1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα 18](#_Toc13807)

[1.1 Βιολογικός - Τεχνητός Νευρώνας 19](#_Toc2685)

[1.2 Μάθηση συναρτήσεων 21](#_Toc12801)

[1.3 Μάθηση με επίβλεψη 24](#_Toc6362)

[1.4 Εφαρμογές νευρωνικών δικτύων 26](#_Toc9555)

[1.5 Σκοπός της εργασίας 29](#_Toc29007)

[2. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα 30](#_Toc23504)

[2.1 Ιστορική Αναδρομή 30](#_Toc2271)

[2.2 Δίκτυα πρόσθιας τροφοδότησης 37](#_Toc29745)

[2.3 Τα δίκτυα Perceptron 39](#_Toc29229)

[2.4 Τα δίκτυα Adaline 42](#_Toc12816)

[2.5 Τα δίκτυα MLP 43](#_Toc6095)

[2.6 Η μέθοδος Back Propagation 45](#_Toc15756)

[2.7 Η μέθοδος Gradient Descent 47](#_Toc12765)

[3. Γενετικοί αλγόριθμοι 48](#_Toc9472)

[3.1 Μέθοδοι κωδικοποίησης 50](#_Toc6467)

[3.2 Γενετικοί τελεστές 51](#_Toc22951)

[3.3 Παράλληλοι γενετικοί αλγόριθμοι 52](#_Toc28260)

[3.4 Εφαρμογές γενετικών αλγορίθμων 53](#_Toc26142)

[4. Μέθοδος - αποτελέσματα 54](#_Toc30222)

[4.1 Τα dataset που χρησιμοποιήθηκαν 54](#_Toc22042)

[4.2 Η προτεινόμενη μέθοδος 55](#_Toc26317)

[4.3 Πειραματικά αποτελέσματα 56](#_Toc16606)

[5. Συμπεράσματα 57](#_Toc29216)

[ΠΑΡΑΡΤΗΜΑ 58](#_Toc13350)

[ΒΙΒΛΙΟΓΡΑΦΙΑ 63](#_Toc28709)

[ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ 63](#_Toc27554)

[ΞΕΝΟΓΛΩΣΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑ 64](#_Toc2474)

# ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1. [Τίτλος]…………………………………………………………….αρ. σελίδας

Πίνακας 2. [Τίτλος]…………………………………………………………….αρ. σελίδας

*[Διαγράψτε αυτή τη σελίδα αν δεν τη χρειάζεστε. Αν έχει εφαρμογή, η αρίθμηση των Πινάκων γίνεται με βάση τον αριθμό του κεφαλαίου που ανήκουν π.χ. Πίνακας 1.1, Πίνακας 1.2, Πίνακας 3.1]*

# ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1.1. Βιολογικός νευρώνας………………………………...…………..…αρ. σελίδας

Εικόνα 1.2. Τεχνητός νευρώνας...………………………………………….....…αρ. Σελίδας

Εικόνα 1.3. Γραφική παράσταση της γραμμικής συνάρτησης...…...............……αρ. Σελίδας

Εικόνα 1.4. Γραφική παράσταση της βηματικής συνάρτησης 0/1...….....………αρ. Σελίδας

Εικόνα 1.5. Γραφική παράσταση της βηματικής συνάρτησης -1/1...…....………αρ. Σελίδας

Εικόνα 1.6. Γραφική παράσταση της συνάρτησης σιγμοειδής συνάρτησης.....…αρ. Σελίδας

Εικόνα 1.7. Γραφική παράσταση της συνάρτησης υπερβολικής εφαπτομένης…αρ. Σελίδας

# ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ

ΤΝΔ..………………………………………………..........…….Τεχνητά Νευρωνικά Δίκτυα

ΤΝ...........................................................................................................Τεχνητή Νοημοσύνη

MLP……………………………………………....Μοντέλο Perceptron πολλών στρωμάτων

# ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ

*[Διαγράψτε αυτή τη σελίδα αν δεν τη χρειάζεστε.]*

# 1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα

Η πρώτη εμφάνιση των ΤΝΔ έγινε τη δεκαετία του 1940 και του 1950 με το βασικό μοντέλο του νευρώνα, που ανέπτυξαν οι Αμερικανοί επιστήμονες McCulloch και Pitts. Την ίδια περίοδο, ο Rosenblatt ανέπτυξε τον πρώτο αλγόριθμο εκπαίδευσης ενός νευρώνα, το γνωστό Perceptron. Ωστόσο, το 1969, οι Minsky και Papert απέδειξαν τις περιορισμένες δυνατότητες του.

Η αναγέννηση των ΤΝΔ πραγματοποιήθηκε τη δεκαετία του 1980 με το μοντέλο του Hopfield και το μοντέλο Perceptron Πολλών Στρωμάτων (Multi-Layer Perceptron ή αλλιώς MLP). Το MLP διαθέτει απεριόριστες δυνατότητες αναπαράστασης συναρτήσεων και διαχωρισμού κλάσεων ειδικά σε συνδυασμό με τον αλγόριθμο εκπαίδευσης Back-Propagation. ‘Ένας αλγόριθμος ικανός για πρώτη φορά να εκπαιδεύσει ένα δίκτυο με περισσότερους από έναν νευρώνες.

Δημιουργήθηκε ένα νέο παράδειγμα υπολογιστικού μοντέλου, το λεγόμενο Connectionist model, με κύριο χαρακτηριστικό την διασύνδεση πολλών απλών υπολογιστικών κόμβων σε δίκτυο και τη δυνατότητα αυτοπροσαρμογής των συνδέσεων του δικτύου. Το μοντέλο αυτό προσέφερε μία νέα προσέγγιση στο πρόβλημα της ΤΝ μέσω της μάθησης, όπου το σύστημα αυτοπροσαρμόζεται χρησιμοποιώντας τα δεδομένα και δεν βασίζεται σε κάποιους προκαθορισμένους κανόνες λογικής για την εξαγωγή συμπερασμάτων.

Σήμερα, υπάρχει μία πληθώρα νευρωνικών μοντέλων που ακολουθούν διάφορα πρότυπα μάθησης, όπως εκπαίδευση με ή χωρίς επίβλεψη, ανταγωνιστικά μοντέλα, αυτο-οργανούμενα μοντέλα, και άλλα. (Διαμαντάρας)

Τα ΤΝΔ είναι μοντέλα που μιμούνται τη λειτουργία των βιολογικών νευρώνων και τη δομή των βιολογικών νευρωνικών δικτύων. Το αντικείμενο των ΤΝΔ είναι η ανάπτυξη και η μελέτη μαθηματικών αλγορίθμων που μιμούνται την αρχιτεκτονική και το πρότυπο των βιολογικών νευρωνικών δικτύων. Το κυριότερο αντικείμενο μελέτης τους είναι η ανάπτυξη αλγορίθμων που θα μιμούνται αυτές τις λειτουργίες όπως αναγνώριση φυσικής γλώσσας, αναγνώριση προσώπων και περιβάλλοντος, πλοήγηση ενός ρομπότ σε περιβάλλον με φυσικά εμπόδια, ανάπτυξη βέλτιστων στρατηγικών για ένα πρόβλημα, εκτέλεση συλλογισμών καταλήγοντας σε λογικά συμπεράσματα, θα έχουν μνήμη και τέλος θα αυτοπροσαρμόζονται σε νέες καταστάσεις και σε γνωστά περιβάλλοντα και θα μαθαίνουν από την εμπειρία τους.

## 1.1 Βιολογικός - Τεχνητός Νευρώνας

Ο νευρώνας είναι ένα μεγάλο σε μέγεθος κύτταρο, και παρά την ύπαρξη αρκετών παραλλαγών του, ανατομικά αποτελείται από τέσσερα βασικά συστατικά. Το σώμα - που αποτελεί και τον πυρήνα του νευρώνα, τους δενδρίτες, τον άξονα και τις συνάψεις. Οι δενδρίτες λειτουργούν ως πύλες εισόδου του νευρώνα, λαμβάνοντας σήματα από γειτονικούς νευρώνες, με την μορφή ηλεκτρικών παλμών. Το σώμα του νευρώνα, επεξεργάζεται αυτά τα εισερχόμενα σήματα με την πάροδο του χρόνου και τα αποστέλλει, σε άλλους νευρώνες μέσω του άξονα. Στόχος των συνάψεων είναι η μετάδοση της ηλεκτρικής δραστηριότητας από τον άξονα - αποστολέα στους δενδρίτες - παραλήπτες, δημιουργώντας έτσι ένα νευρωνικό δίκτυο. Κατά τη μετάδοση αυτής της ηλεκτρικής δραστηριότητας, το ποσοστό που μεταδίδεται τελικά ονομάζεται συναπτικό βάρος. Κάθε νευρώνας συλλέγει όλο το ηλεκτρικό φορτίο που δέχεται από κάθε σύναψη στους δενδρίτες του, ζυγίζοντας το εισερχόμενο φορτίο με το αντίστοιχο συναπτικό βάρος. Όσο πιο ισχυρή είναι η συναπτική ζεύξη τόσο πιο έντονα συμμετέχει το συγκεκριμένο φορτίο εισόδου στο συνολικό άθροισμα. Αν το άθροισμα του φορτίου ξεπερνάει κάποιο κατώφλι τότε ο άξονας αρχίζει να παράγει ηλεκτρικούς παλμούς με μεγάλη συχνότητα, οπότε λέμε ότι ο νευρώνας πυροβολεί και θεωρούμε τη σύναψη ενισχυτική. Αλλιώς ο νευρώνας παράγει πολύ αραιά παλμούς σε τυχαίες στιγμές οπότε λέμε ότι ο νευρώνας είναι αδρανής και έτσι θεωρούμε τη σύναψη ως ανασταλτική.[[1]](#footnote-0)



(Εικόνα 1.1) Αναπαράσταση βιολογικού νευρώνα (Βλαχάβας et al., 2006)

Ο εγκέφαλος είναι σε θέση να λαμβάνει πολύπλοκες αποφάσεις εκπληκτικά γρήγορα. Αυτό είναι αποτέλεσμα στο γεγονός πως η υπολογιστική ικανότητα του εγκεφάλου καθώς και η πληροφορία που αυτός περιέχει είναι διαμοιρασμένα σε όλο του τον όγκο. Αυτά τα χαρακτηριστικά αποτελούν και το μεγαλύτερο κίνητρο έτσι ώστε τα ΤΝΔ να μοντελοποιήσουν τον εγκέφαλο. Ο τεχνητός νευρώνας είναι ένα υπολογιστικό μοντέλο τα μέρη του οποίου μπορούν να αντιστοιχιστούν άμεσα με αυτά του βιολογικού νευρώνα. Δέχεται κάποια σήματα εισόδου, τα οποία σε αντίθεση με τους ηλεκτρικούς παλμούς του εγκεφάλου, αντιστοιχούν σε συνεχείς μεταβλητές. Κάθε σήμα εισόδου, μεταβάλλεται από μία τιμή βάρους, που αντιστοιχεί στο ρόλο της σύναψης σε έναν βιολογικό νευρώνα. Σε περίπτωση επιταχυντικής λειτουργίας της σύναψης, αυτή η τιμή βάρους μπορεί να είναι θετική ενώ σε περίπτωση επιβραδυντικής λειτουργίας μπορεί να είναι αρνητική. Το σώμα του τεχνητού νευρώνα χωρίζεται σε δύο τμήματα. Το πρώτο αποτελείται από τον αθροιστή, ο οποίος προσθέτει τα επηρεασμένα από τα βάρη σήματα εισόδου. Το δεύτερο τμήμα αποτελείται από τη συνάρτηση ενεργοποίησης, ένα είδος φίλτρου που διαμορφώνει την τελική τιμή του σήματος εξόδου. Διευκρινίζεται ότι η μοναδικότητα της εξόδου του νευρώνα έχει να κάνει με την τιμή εξόδου και όχι με το πόσες έξοδοι υπάρχουν. Μπορεί δηλαδή ένας νευρώνας να έχει πολλές εξόδους, όλες όμως θα έχουν την ίδια τιμή. Υπάρχουν αρκετές περιπτώσεις, συνήθως είναι οι περισσότερες, κατά τις οποίες θα θεωρούμε πως υπάρχει και ένα επιπλέον βάρος, το οποίο χαρακτηρίζεται ως πόλωση (bias) ή αλλιώς παράγοντας προδιάθεσης του νευρώνα. Η μόνη διαφορά του από τα υπόλοιπα είναι ότι επιδρά συνεχώς σε μία τιμή εισόδου με τιμή ένα (1). Ο όρος αυτός πρόκειται για εξωτερικό ερέθισμα το οποίο προστίθεται με τα εισερχόμενα σήματα, επομένως δεν πρέπει να αποδίδεται στο εσωτερικό του νευρώνα. (Βλαχάβας et al., 2006)



(Εικόνα 1.2) Τεχνητός νευρώνας (Suzuki, 2011)

## 1.2 Μάθηση συναρτήσεων

Όπως αναφέρθηκε στην προηγούμενη ενότητα, η συνάρτηση ενεργοποίησης αποτελεί ένα από τα δύο βασικά τμήματα του τεχνητού νευρώνα, διότι ορίζει την έξοδο του. Υπάρχουν αρκετά είδη συναρτήσεων που μπορούν να χρησιμοποιηθούν, με πιο δημοφιλή τις παρακάτω.



Γραμμική συνάρτηση:

C:/Users/Evita/AppData/Local/Temp/wps.jMZJNzwps

(Εικόνα 1.3) Γραφική παράσταση της γραμμικής συνάρτησης

Βηματική συνάρτηση:



(Εικόνα 1.4) Γραφική παράσταση της βηματικής συνάρτησης 0/1

Περισσότερο χρησιμοποιείται σε θεωρητικές μελέτες, και είναι η πιο απλή συνάρτηση ενεργοποίησης. (Chiarandini & Τσούλος)

Αλλά και με την μορφή

(Εικόνα 1.5) Γραφική παράσταση της βηματικής συνάρτησης -1/1

Σιγμοειδής συνάρτηση:



(Εικόνα 1.6) Γραφική παράσταση της σιγμοειδής συνάρτησης

Χρησιμοποιείται περισσότερο σε πρακτικές εφαρμογές. (Chiarandini) Η γραφική της παράσταση έχει το σχήμα “S” και είναι η πιο κοινή μορφή συνάρτησης ενεργοποίησης που χρησιμοποιείται για την κατασκευή ΤΝΔ. (*Haykin, 2010)*



Υπερβολική εφαπτομένη:

(Εικόνα 1.7) Γραφική παράσταση της συνάρτησης υπερβολικής εφαπτομένης

## 1.3 Μάθηση με επίβλεψη

Τα ΤΝΔ βασίζονται στη λειτουργία του ανθρώπινου εγκεφάλου. Έτσι, η πρόκληση που αντιμετωπίζουν είναι η εύρεση κατάλληλων αλγορίθμων εκπαίδευσης των δικτύων καθώς και ανάκλησης της πληροφορίας που περιέχεται. Στόχος των αλγορίθμων αυτών είναι η προσομοίωση ευφυών διαδικασιών όπως αυτή της αναγνώρισης προτύπων, της λήψης αποφάσεων καθώς και της επίλυσης προβλημάτων. Με σκοπό την επίτευξη των παραπάνω, απαιτείται ο ορισμός του κατάλληλου περιβάλλοντος εκπαίδευσης, μία διαδικασία που μπορεί να κατηγοριοποιηθεί ανάλογα με τον τρόπο λειτουργίας της. Οι πιο βασικές κατηγορίες είναι αυτή της μάθησης με εκπαιδευτή και αυτή χωρίς εκπαιδευτή.

Σε αυτή την πτυχιακή εργασία, θα εστιάσουμε στη μάθηση με εκπαιδευτή διότι χρησιμοποιείται στον μεγαλύτερο όγκο των δικτύων. Η μάθηση με εκπαιδευτή αναφέρεται επίσης ως μάθηση με επίβλεψη ή αλλιώς επιβλεπόμενη μάθηση. Σε αυτή την κατηγορία μάθησης, το περιβάλλον είναι άγνωστο στο νευρωνικό δίκτυο. Ωστόσο, μπορούμε να θεωρήσουμε ότι ο εκπαιδευτής έχει γνώση του περιβάλλοντος, η οποία αντιπροσωπεύεται από ένα δείγμα προτύπων έχοντας αντιστοίχηση μεταξύ ενός σήματος εισόδου και της αντίστοιχης επιθυμητής απόκρισης. Αυτή η πληροφορία είναι γνωστή και ως a-priori (εκ των προτέρων). Αξίζει να σημειωθεί πως η εργασία της συλλογής αυτών των προτύπων είναι χρονοβόρα και ακριβή, ειδικά όταν αντιμετωπίζουμε μεγάλης κλίμακας προβλήματα μάθησης.

Όπως αναφέρθηκε, σε αυτή την κατηγορία μάθησης υπάρχει αντιστοίχιση ανάμεσα στην είσοδο και στην έξοδο. Έτσι, το δίκτυο συνέχεια επεξεργάζεται τις εισόδους και τις συγκρίνει με το επιθυμητό αποτέλεσμα της εξόδου. Στις περιπτώσεις κατά τις οποίες δεν υπάρχει αντιστοίχηση ανάμεσα στο πραγματικό και το επιθυμητό αποτέλεσμα της εξόδου, το σύστημα αναπροσαρμόζει τα βάρη του. Αυτή είναι μία επαναλαμβανόμενη διαδικασία, κάνοντας αρκετές φορές επεξεργασία το σύνολο των προτύπων καθώς τα βάρη βελτιώνονται συνεχώς.

Πρέπει να σημειωθεί βέβαια ότι κάποια δίκτυα δεν μαθαίνουν ποτέ. Ένας βασικός λόγος που μπορεί να προκύψει ένα τέτοιο αποτέλεσμα, είναι από λανθασμένα σύνολα δεδομένων ή μη επαρκείς πληροφορίες. Όσο περισσότερα δεδομένα έχουμε στη διάθεση μας, τόσο το καλύτερο διότι αρκετά δεδομένα μπορούν να κρατηθούν για δοκιμή. (Anderson & McNeill, 1992)

Σύμφωνα με τον Διαμαντάρα (2007), οι κύριοι αλγόριθμοι εκπαίδευσης με επίβλεψη περιλαμβάνουν τα δίκτυα Perceptron, Adaline, Back Propagation καθώς και τα αναδρομικά Back Propagation. Παράλληλα οι Βλαχάβας et al. (2006) αναφέρουν ότι στη μάθηση με επίβλεψη διακρίνονται δύο είδη προβλημάτων. Το πρώτο είδος αφορά τα προβλήματα ταξινόμησης (classification), τα οποία αφορούν τη δημιουργία μοντέλων πρόβλεψης διακριτών κλάσεων όπως για παράδειγμα η ομάδα αίματος. Το δεύτερο είδος αφορά τα προβλήματα παρεμβολής (regression), τα οποία αφορούν τη δημιουργία μοντέλων πρόβλεψης αριθμητικών τιμών. Οι κυριότερες τεχνικές μηχανικής μάθησης με επίβλεψη είναι η μάθηση εννοιών, τα δένδρα ταξινόμησης ή απόφασης, η μάθηση κανόνων, η μάθηση κατά περίπτωση, η μάθηση κατά Bayes, η γραμμική παρεμβολή, τα νευρωνικά δίκτυα καθώς και οι μηχανές διανυσμάτων υποστήριξης (Support Vector Machines). [[2]](#footnote-1)

## 1.4 Εφαρμογές νευρωνικών δικτύων

Τα ΤΝΔ είναι ιδιαίτερα δημοφιλή σε προβλήματα που περιέχουν μη-προβλέψιμες λειτουργίες και τα οποία δεν είναι πλήρως κατανοητά. Τέτοιου είδους προβλήματα υπάρχουν σε πολλές ανθρώπινες δραστηριότητες που σχετίζονται με κατηγοριοποίηση (classification), αναγνώριση (recognition/identification), αποτίμηση (assessment), και πρόβλεψη (forecasting/prediction).

Σχετικά με τα προβλήματα κατηγοριοποίησης, μπορούμε να αναφέρουμε αρχικά τον ιατρικό τομέα. Αξίζει να αναφερθεί η διάγνωση παθήσεων, η βιοπληροφορική, η ανάπτυξη νέων φαρμάκων για ασθένειες καθώς και ανίχνευση βλαβών σε μηχανήματα. Μέσα από αυτές τις εφαρμογές επιταχύνεται σημαντικά η χρονοβόρα διαδικασία ελέγχου των δεδομένων ιατρικών εξετάσεων από ιατρούς, χωρίς όμως κάτι τέτοιο να σημαίνει ότι οι ιατροί δεν είναι απαραίτητοι. Στον τομέα της άμυνας, γίνεται η κατηγοριοποίηση εικόνων προερχόμενων από συσκευές radar, sonar και άλλα. Στις επιχειρήσεις, γίνεται κατηγοριοποίηση των πελατών βάση των αγοραστικών τους συνηθειών.

Σε προβλήματα αναγνώρισης, αξίζει να επισημάνουμε τον τραπεζικό τομέα στον οποίο γίνεται έλεγχος γνησιότητας των υπογραφών αλλά και των τραπεζογραμματίων. Στον τομέα της πληροφορικής έχουμε την αναγνώριση ήχου, εικόνας καθώς και γραπτού κειμένου είτε αυτό είναι χειρόγραφο είτε τυπωμένο. Οι εφαρμογές OCR, δηλαδή οπτική αναγνώρισης χαρακτήρων είναι από τις πιο διάσημες εφαρμογές ΤΝΔ. Τέλος μπορούμε να συμπεριλάβουμε και την ανάπτυξη επιχειρημάτων καθώς χρειάζεται αποτίμηση των δεδομένων για την υποστήριξη ενός επιχειρήματος.

Σε προβλήματα αποτίμησης, έχουμε την παρακολούθηση στόχων στον τομέα της άμυνας, τον εντοπισμό κίνησης, ταύτιση δακτυλικών αποτυπωμάτων καθώς και ανάλυση εικόνας σε συστήματα επιτήρησης στον τομέα της ασφάλειας.Επιπλέον, στη μηχανολογία γίνεται παρακολούθηση, επιθεώρηση και έλεγχος προϊόντων.

Τέλος, στον τομέα της πρόβλεψης, έχουμε την πρόβλεψη ισοτιμίας νομισμάτων και τιμών μετοχών, καθώς και την πρόβλεψη πωλήσεων. Στη γεωργία γίνεται πρόβλεψη παραγωγής κυρίως με χρήση δορυφορικών εικόνων. Στη μετεωρολογία έχουμε πρόβλεψη καιρού. Επίσης μπορούμε να συμπεριλάβουμε την αυτόματη πλοήγηση στο χώρο καθώς απαιτεί πρόβλεψη της κίνησης στο χώρο και δυναμικών εμποδίων. Τέλος, η λήψη αποφάσεων διότι περιλαμβάνει την πρόβλεψη των συνεπειών διάφορων επιλογών.

Τα τελευταία χρόνια τα ΤΝΔ χρησιμοποιούνται σε συστήματα ελέγχου που βασίζονται στην ασαφή λογική (neuofuzzy systems) με κύριο ρόλο τον υπολογισμό της συνάρτησης συγγένειας.

Σε οικονομικά θέματα όπως χρηματιστηριακές προβλέψεις,

η αναγνώριση εικόνων καθώς και φωνής, η μνήμη, η κατανόηση και η παραγωγή της γλώσσας, Επιπλέον, η κατάστρωση στρατηγικής και η επιλογή της καλύτερης βάσει διάφορων κριτηρίων κόστους, η λογική.

Αξίζει να σημειωθεί η μάθηση και η αυτοπροσαρμογή σε νέο περιβάλλον καθώς και καταστάσεις. Ιδιαίτερη σημασία έχει η εφαρμογή των γενετικών αλγορίθμων τόσο στην οικονομία όσο και στο εμπόριο προβλέποντας οικονομικά μεγέθη, ανιχνεύοντας απάτες σε ηλεκτρονικές συναλλαγές καθώς και εκτιμώντας την αξία ακινήτων. Επίσης στη βιομηχανία μέσω βιομηχανικών ελέγχων, ρύθμιση ηλεκτρικού φορτίου, ρομποτική και εφαρμογές σε οχήματα. Τέλος, η αναγνώριση προτύπων, ασφάλεια, ηλεκτρονικά παιχνίδια αφορούν εφαρμογές των γενετικών αλγορίθμων.

(Τσούλος, Βλαχάβας)

*(Βλαχάβας)*

Πρέπει να τονιστεί ότι αν και ένα εκπαιδευόμενο ΤΝΔ μπορεί να αναγνωρίσει δεδομένα τα οποία δεν έχει δει ποτέ του, αυτό δεν συμβαίνει στην περίπτωση που τα δεδομένα δεν ανήκουν στην ίδια κατηγορία προβλημάτων για την οποία έχει εκπαιδευτεί. Δεν υπάρχουν ΤΝΔ γενικού σκοπού τα οποία μπορούν να αντιμετωπίζουν διάφορα ετερογενή προβλήματα. Υπάρχουν όμως προγραμματιστικά περιβάλλοντα τα οποία επιτρέπουν τη δόμηση ΤΝΔ για διάφορους σκοπούς. Αυτό είναι πολύ σημαντικό καθώς η απάντηση στο ερώτημα “ποιο είδος ΤΝΔ είναι κατάλληλο για δεδομένο πρόβλημα” δεν είναι ούτε εύκολη, ούτε μοναδική. Συνήθως κάθε πρόβλημα είναι ειδική περίπτωση και η αντιμετώπιση του με ΤΝΔ απαιτεί μελέτη και πειραματισμό.

## 1.5 Σκοπός της εργασίας

Αυτή η πτυχιακή αποσκοπεί ..

# Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα

## Ιστορική Αναδρομή

Τη δεκαετία του 1940 υπήρξε αρκετά έντονη δραστηριότητα προς την κατεύθυνση της μελέτης των βιολογικών νευρωνικών δικτύων και της μαθηματικής μοντελοποίησης τους. Πρωτοπόροι στον τομέα αυτό ήταν οι Αμερικανοί επιστήμονες McCulloch και Pitts, οι οποίοι περιέγραψαν ένα απλό μοντέλο της δραστηριότητας του νευρώνα όπου η κατάσταση του περιγράφεται από ένα δυαδικό αριθμό. Όταν αυτός ο αριθμός έχει την τιμή 0 σημαίνει πως ο νευρώνας είναι αδρανής (δε πυροβολεί) ενώ όταν έχει την τιμή 1 σημαίνει πως ο νευρώνας πυροβολεί στη μέγιστη ταχύτητα.

(Βλαχάβας)

Εχουν το μεγάλο πλεονέκτημα της ανοχής που παρουσιάζουν σε δεδομένα εκπαίδευσης με θόρυβο, δηλαδή δεδομένα που περιστασιακά έχουν λανθασμένες τιμές (πχ λάθη καταχώρησης). Από την άλλη όμως αδυνατούν να εξηγήσουν ποιοτικά τη γνώση που μοντελοποιούν.

Τα Τ.Ν.Δ. (artificial neural networks) είναι συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος τεχνητών νευρώνων οργανωμένων σε δομές παρόμοιες με αυτές του ανθρώπινου εγκεφάλου. Συνήθως οι τεχνητοί νευρώνες είναι οργανωμένοι σε μία σειρά από στρώματα ή επίπεδα (layers). Το πρώτο από αυτά τα επίπεδα ονομάζεται επίπεδο εισόδου (input layer) και χρησιμοποιείται για την εισαγωγή δεδομένων. Τα στοιχεία του δηλαδή δεν είναι ουσιαστικά νευρώνες, γιατί δεν εκτελούν κάποιον υπολογισμό (δεν έχουν βάρη εισόδου, ούτε συνάρτηση ενεργοποίησης). Στη συνέχεια, μπορεί να ακολουθούν προαιρετικά, ένα ή περισσότερα ενδιάμεσα ή κρυφά επίπεδα (hidden layers), ενώ στο τέλος υπάρχει το επίπεδο εξόδου (output layer).

Ένα παράδειγμα Τ.Ν.Δ είναι το 3-4-2 το οποίο φαίνεται να έχει 1 επίπεδο εισόδου με 3 εισόδους, 4 κρυφούς νευρώνες και 1 επίπεδο εξόδου με 2 εξόδους.

Οι νευρώνες στα ΤΝΔ μπορεί να είναι πλήρως ή μερικώς συνδεδεμένοι. Πλήρως συνδεδεμένοι (fully connected) είναι εκείνοι οι οποίοι συνδέονται με όλους τους υπόλοιπους νευρώνες. Σε κάθε άλλη περίπτωση οι νευρώνες είναι μερικώς συνδεδεμένοι (partially connected). Μία συνήθης περίπτωση μερικής διασύνδεσης είναι αυτή στην οποία οι νευρώνες ενός επιπέδου είναι πλήρως διασυνδεδεμένοι με αυτούς του επόμενου επιπέδου. Όταν δεν υπάρχουν συνδέσεις μεταξύ νευρώνων ενός επιπέδου και νευρώνων προηγούμενου επιπέδου (όταν δηλ. η ροή πληροφορίας είναι πρόσθιας κατεύθυνσης) τα ΤΝΔ χαρακτηρίζονται ως δίκτυα με πρόσθια τροφοδότηση (feedforward). Στην αντίθετη περίπτωση, καθώς και στην περίπτωση συνδέσεων μεταξύ νευρώνων ίδιου επιπέδου, τα ΤΝΔ χαρακτηρίζονται ως δίκτυα με ανατροφοδότηση (feedback ή recurrent). Αν και σε ορισμένες περιπτώσεις τα δίκτυα με ανατροφοδότηση είναι πολύ χρήσιμα, στην πλειοψηφία των εφαρμογών νευρωνικών δικτύων χρησιμοποιούνται δίκτυα πρόσθιας τροφοδότησης.

Υπάρχουν 4 ιδιότητες που είναι άρηκτα συνδεδεμένες με τα ΤΝΔ.

1.Η ικανότητα τους να μαθαίνουν μέσω παραδειγμάτων (learn by example)

2.Η δυνατότητα θεώρησης τους ως κατανεμημένη μνήμη (distributed memory) και ως μνήμη συσχέτισης (associative memory).

3.Η μεγάλη τους ανοχή σε σφάλματα (fault-tolerant).

4.Η ικανότητα τους για αναγνώριση προτύπων (pattern recognition).

Αν και τα ΤΝΔ δεν είναι τα μόνα συστήματα με ικανότητα μάθησης μέσω παραδειγμάτων, εντούτοις διακρίνονται για την ικανότητα τους να οργανώνουν την πληροφορία των δεδομένων εισόδου σε χρήσιμες μορφές. Αυτές οι μορφές αποτελούν στην ουσία ένα μοντέλο που αναπαριστά τη σχέση που ισχύει μεταξύ των δεδομένων εισόδου και εξόδου.

Ο χαρακτηρισμός των ΤΝΔ ως κατανεμημένη μνήμη, πηγάζει από το ότι η κωδικοποίηση που δημιουργούν είναι κατανεμημένη σε όλα τα βάρη της συνδεσμολογίας τους. Για τον ίδιο λόγο τα ΤΝΔ χαρακτηρίζονται και ως μνήμες συσχέτισης. Μία μνήμη συσχέτισης δεν αποθηκεύει πληροφορία με τον παραδοσιακό τρόπο αλλά μέσω κατάλληλων συσχετίσεων που δημιουργεί από τα δεδομένα εκπαίδευσης. Η ανάκληση της πληροφορίας γίνεται με βάση το περιεχόμενο και όχι τη διεύθυνση, όπως δηλαδή συμβαίνει και με τον ανθρώπινο εγκέφαλο. Η παραπάνω οργάνωση, κάνει ορισμένα είδη ΤΝΔ να είναι πολύ ανεκτικά σε μικρές αλλαγές στα σήματα εισόδου, δηλ. Είναι σε θέση να παράγουν τη σωστή έξοδο ακόμη και αν τα δεδομένα εισόδου είναι λίγο διαφορετικά (για παράδειγμα λόγω θορύβου) ή και ελλιπή.

Τα ΤΝΔ, όπως και τα βιολογικά, έχουν μεγάλη ανοχή σε δομικά σφάλματα. Αυτό σημαίνει ότι η κακή λειτουργία ή η καταστροφή ενός νευρώνα ή κάποιων συνδέσεων δεν είναι ικανή να διαταράξει σημαντικά τη λειτουργία τους καθώς, όπως αναφέρθηκε, η πληροφορία που εσωκλείουν δεν είναι εντοπισμένη σε συγκεκριμένο σημείο αλλά διάχυτη σε όλο το δίκτυο. Γενικά, το μέγεθος του σφάλματος λόγω “δομικών αστοχιών” είναι ανάλογο του ποσοστού των κατεστραμμένων συνδέσεων.

Τέλος, τα ΤΝΔ έχουν εξαιρετική ικανότητα αναγνώρισης προτύπων καθώς δεν επηρεάζονται από ελλιπή ή/και με θόρυβο δεδομένα. Από τη στιγμή που ένα ΤΝΔ εκπαιδευτεί στο να αναγνωρίζει συνθήκες και καταστάσεις, απαιτείται ένας μόνο κύκλος λειτουργίας τους για να προσδιορίσουν μία συγκεκριμένη κατάσταση.

Οι τελευταίες δύο ιδιότητες κάνουν τα ΤΝΔ ιδανικά για χρήση σε αυτοματισμούς που θα λειτουργήσουν σε αντίξοες συνθήκες όπως για παράδειγμα σε διαστημικές αποστολές, σε χώρους με ραδιενέργεια και σε πεδία μάχης.

Εμπνευσμένη από τη δομή και τη λειτουργία του ανθρώπινου εγκεφάλου είναι και η έρευνα σχετικά με τα Τεχνητά Νευρωνικά Δίκτυα (Τ.Ν.Δ.) Το πιο βασικό δομικό στοιχείο του εγκεφάλου είναι οι νευρώνες, δηλαδή τα νευρικά κύτταρα τα οποία δημιουργούν ένα πυκνό δίκτυο επικοινωνίας μεταξύ τους. Κίνητρο για τη μελέτη του νευρώνα και των νευρωνικών δικτύων είναι η ελπίδα ανακάλυψης ενός νέου υπολογιστικού μοντέλου βασισμένου σε μία δικτυακή δομή παρόμοια με αυτού του εγκεφάλου. (Βλαχάβας et al., 2006)

Σύμφωνα με τον Haykin ένα ν.δ. ως προσαρμόσιμη μηχανή είναι ένας τεράστιος παράλληλος επεξεργαστής με κατανεμημένη αρχιτεκτονική, ο οποίος αποτελείται από απλές μονάδες επεξεργασίας και έχει από τη φύση του τη δυνατότητα να αποθηκεύει εμπειρική γνώση και να την καθιστά διαθέσιμη για χρήση.

Υπάρχουν δύο κοινά σημεία με τον ανθρώπινο εγκέφαλο. Αρχικά, το δίκτυο προσλαμβάνει τη γνώση από το περιβάλλον του, μέσω μίας διαδικασίας μάθησης. Επίσης, η ισχύς των συνδέσεων μεταξύ των νευρώνων, που αποκαλείται αλγόριθμος μάθησης και η λειτουργία του είναι να τροποποιεί τα συναπτικά βάρη του δικτύου με τον κατάλληλο τρόπο για την επίτευξη του επιθυμητού στόχου.

Είναι προφανές ότι ένα ν.δ. οφείλει την υπολογιστική ισχύ του κατά πρώτον στην παράλληλη, κατανεμημένη δομή του και κατά δεύτερον στην ικανότητα του να μαθαίνει, και ως εκ τούτου να γενικεύει. Λόγω της μάθησης και της γενίκευσης τα ν.δ. έχουν την δυνατότητα να βρίσκουν καλές προσεγγιστικές λύσεις σε πολύπλοκα (μεγάλης κλίμακας) προβλήματα, τα οποία είναι μη επιδεκτικά σε λύσεις. Ωστόσο, τα ν.δ. δεν μπορούν να παρέχουν λύση αν λειτουργούν ατομικά. Συγκεκριμένα, ένα πολύπλοκο πρόβλημα αποσυντίθεται σε έναν αριθμό σχετικά απλών εργασιών και τα νευρωνικά δίκτυα αναλαμβάνουν ένα υποσύνολο των εργασιών που ταιριάζουν με τις εγγενείς δυνατότητες τους. Θα πρέπει ωστόσο, να αποδεχτούμε ότι έχουμε να διανύσουμε πολύ δρόμο ακόμη μέχρι να μπορέσουμε να κατασκευάσουμε μία αρχιτεκτονική υπολογιστών που θα μιμείται τον ανθρώπινο εγκέφαλο.

Μερικές από τις πιο χρήσιμες ιδιότητες και δυνατότητες που μας προσφέρουν τα τ.ν.δ. είναι η μη γραμμικότητα, η αντιστοίχιση εισόδου - εξόδου κυρίως για τη μάθηση που αναφέρουμε ως επιβλεπόμενη. Όσο πιο προσαρμοστικό κάνουμε ένα σύστημα, διασφαλίζοντας ταυτόχρονα ότι παραμένει διαρκώς σταθερό, τόσο πιο εύρωστο θα είναι και τόσο καλύτερα θα αποδίδει όταν θα κληθεί να λειτουργήσει σε ένα μη σταθερό περιβάλλον. Στο πλαίσιο της ταξινόμησης προτύπων ένα Ν.Δ. μπορεί να σχεδιαστεί ώστε να παρέχει πληροφορία όχι μόνο για το ποιο συγκεκριμένο πρότυπο θα επιλεγεί, αλλά επίσης σχετικά με τον βαθμό εμπιστοσύνης στην ληφθείσα απόφαση, κάτι το οποίο μπορεί να χρησιμοποιηθεί για την απόρριψη των διφορούμενων μοτίβων, εάν προκύψουν και κατ’ επέκταση τη βελτίωση της απόδοσης του δικτύου. Επίσης, το ν.δ. μπορεί να χειριστεί με φυσικό τρόπο τη σχετική με το περιεχόμενο πληροφορία (contextual information) καθώς η γνώση αντιπροσωπεύεται από την ίδια τη δομή και κατάσταση ενεργοποίησης του. Κάθε νευρώνας στο δίκτυο ενδεχομένως να επηρεάζεται από τη συνολική δραστηριότητα όλων των άλλων νευρώνων του δικτύου. Επιπλέον, λόγω της κατανεμημένης φύσης της πληροφορίας που αποθηκεύεται στο δίκτυο, ένα ν.δ. υλοποιημένο σε μορφή hardware, έχει την εγγενή δυνατότητα να είναι ανεκτικό σε βλάβες, ή εύρωστο υπό την έννοια ότι η απόδοση του μειώνεται βαθμιαία και ομαλά υπό αντίξοες συνθήκες λειτουργίας. Ένα ακόμη χαρακτηριστικό είναι πως η μαζικά παράλληλη φύση ενός ν.δ. το καθιστά ενδεχομένως γρήγορο για τον υπολογισμό συγκεκριμένων εργασιών, κάτι το οποίο το κάνει να είναι ιδιαίτερα κατάλληλο για χρήση τεχνολογίας πολύ μεγάλης κλίμακας ολοκλήρωσης (VLSI).Τέλος, η σχεδίαση ενός ν.δ. δανείζεται στοιχεία από τη λειτουργία του ανθρώπινου εγκεφάλου, ο οποίος είναι η ζωντανή απόδειξη ότι η εύρωστη, παράλληλη επεξεργασία δεν είναι μόνο φυσικά εφικτή, αλλά επίσης γρήγορη και ισχυρή.

Μεταξύ των δεκαετιών 1940 και 1950, έκαναν την εμφάνιση τους και τα πρώτα μοντέλα των Τεχνητών Νευρωνικών Δικτύων (Τ.Ν.Δ.), ξεκινώντας από το βασικό μοντέλο του νευρώνα των Αμερικανών επιστημόνων McCulloch-Pitts και τον πρώτο αλγόριθμο εκπαίδευσης ενός νευρώνα, το γνωστό Perceptron του Frank Rosenblatt. To 1969 όμως αποδείχτηκε από τους Minsky και Papert πως αυτός ο αλγόριθμος είχε περιορισμένες δυνατότητες. To 1980 αναπτύχθηκε το μοντέλο του Hopfield και το μοντέλο Perceptron πολλών στρωμάτων (Multi-Layer Perceptron ή MLP) σε συνδυασμό με τον αλγόριθμο εκπαίδευσης Back-Propagation. Για πρώτη φορά, εμφανίζεται ένας αλγόριθμος ικανός να εκπαιδεύσει ένα δίκτυο με περισσότερους από έναν νευρώνες. Ένα νέο υπολογιστικό μοντέλο εμφανίζεται, το οποίο προσφέρει μία νέα προσέγγιση στο πρόβλημα της Τεχνητής Νοημοσύνης μέσω της μάθησης, το λεγόμενο Connectionist model, με κύριο χαρακτηριστικό την διασύνδεση πολλών απλών υπολογιστικών κόμβων σε δίκτυο και τη δυνατότητα αυτοπροσαρμογής των συνδέσεων του δικτύου χρησιμοποιώντας δεδομένα χωρίς να βασίζεται σε κάποιους προκαθορισμένους κανόνες λογικής για την εξαγωγή συμπερασμάτων. Αυτή η καινούργια υπολογιστική πλατφόρμα, θα είναι πιο κατάλληλη για ανάπτυξη ευφυών αλγορίθμων και γενικότερα διαδικασιών σχετιζόμενων με τη νοημοσύνη όπως η μάθηση, η μνήμη, η γενίκευση, καθώς και η ομαδοποίηση προτύπων.[ Haykin & Διαμαντάρας]

Ο ανθρώπινος εγκέφαλος είναι ένας εξαιρετικά πολύπλοκος, μη γραμμικός, παράλληλος υπολογιστής. Θα μπορούσαμε να τον αναφέρουμε και ως ένα σύστημα επεξεργασίας πληροφορίας. Έχει τη δυνατότητα να οργανώνει τα δομικά του στοιχεία, γνωστά ως νευρώνες με τρόπο ώστε να εκτελούν συγκεκριμένους υπολογισμούς όπως είναι η αναγνώριση προτύπων καθώς και η αντίληψη και ο έλεγχος της κίνησης [2]

Ένα χαρακτηριστικό παράδειγμα είναι η ανθρώπινη όραση, η οποία είναι μία διαδικασία επεξεργασίας πληροφοριών. Είναι ευθύνη του οπτικού συστήματος να μας παρέχει μία αναπαράσταση του περιβάλλοντος μας, και ακόμη πιο σημαντικό, να μας προμηθεύει με τις πληροφορίες που χρειαζόμαστε για να επικοινωνήσουμε μ’ αυτό. Συγκεκριμένα, ο εγκέφαλος εκτελεί διαρκώς και ασταμάτητα εργασίες αναγνώρισης που βασίζονται στην αντίληψη όπως είναι η αναγνώριση ενός οικείου προσώπου που βρίσκεται σε μία άγνωστη σκηνή, την ίδια στιγμή που εργασίες πολύ μικρότερου βαθμού πολυπλοκότητας απαιτούν πολύ μεγαλύτερους χρόνους για να εκτελεστούν από έναν ισχυρό υπολογιστή.

Το έργο στο επιστημονικό πεδίο των Τ.Ν.Δ. ( χάριν συντομίας, αποκαλούνται συνήθως <<νευρωνικά δίκτυα>>) βασίστηκε, από τις απαρχές του, στο γεγονός ότι ο ανθρώπινος εγκέφαλος εκτελεί τους υπολογισμούς με εντελώς διαφορετικό τρόπο από το συμβατικό ψηφιακό υπολογιστή. Στην πλέον γενική μορφή του, ένα νευρωνικό δίκτυο είναι μία μηχανή σχεδιασμένη ώστε να μοντελοποιεί τον τρόπο με τον οποίο ο εγκέφαλος εκτελεί μία συγκεκριμένη εργασία ή λειτουργία. Για να επιτυγχάνουν καλή απόδοση τα νευρωνικά δίκτυα χρησιμοποιούν τεράστιο αριθμό απλών, διασυνδεδεμένων μεταξύ τους υπολογιστικών κυττάρων, τα οποία αποκαλούνται νευρώνες ή μονάδες επεξεργασίας.

## 2.2 Δίκτυα πρόσθιας τροφοδότησης

Τα ΤΝΔ πρόσθιας τροφοδότησης (feedforward) είναι η πιο απλή μορφή νευρωνικών δικτύων και το όνομα τους οφείλεται στο ότι η ροή της πληροφορίας μέσα στο δίκτυο είναι μονής κατεύθυνσης. Σε αυτά υπάρχει ένα επίπεδο εισόδου, ένα επίπεδο εξόδου και προαιρετικά, ένα ή περισσότερα ενδιάμεσα, κρυφά επίπεδα.

Δύο είναι τα θέματα τα οποία ανακύπτουν στην υλοποίηση των νευρωνικών δικτύων του τύπου αυτού. Το πρώτο αφορά στη μάθηση, δηλαδή τον τρόπο με τον οποίο το δίκτυο εκπαιδεύεται για να έχει την επιθυμητή συμπεριφορά. Στα δίκτυα πρόσθιας τροφοδότησης χρησιμοποιούνται μέθοδοι μάθησης με επίβλεψη.

Το δεύτερο θέμα αφορά στην τοπολογία του δικτύου, δηλαδή το πόσα κρυφά επίπεδα θα έχει το δίκτυο, από πόσους νευρώνες θα αποτελείται το κάθε επίπεδο καθώς και πως θα συνδέονται οι νευρώνες μεταξύ τους. Δεν υπάρχει κανόνας για τον προσδιορισμό αυτών των μεγεθών. Για τον προσδιορισμό του αριθμού των νευρώνων στα επίπεδα εισόδου και εξόδου το πρόβλημα είναι σχετικά ευκολότερο καθώς τα δεδομένα του προβλήματος μπορούν να βοηθήσουν. Εάν λόγου χάρη πρέπει να αναγνωριστούν τα ψηφία 0 ως 9 σε γραπτή απεικόνιση και η αναπαράσταση τους γίνει με μήτρα 6x4 εικονοστοιχίων (pixels), τότε χρειάζονται 6x4=24 νευρώνες εισόδου, με δυνατές τιμές εισόδου 0 και 1. Αν χρησιμοποιηθεί κωδικοποίηση 4 bit για τα δέκα ψηφία 0 εως 9, τότε χρειάζονται 4 νευρώνες στο επίπεδο εξόδου. Εναλλακτικά, μπορεί να χρησιμοποιηθούν 10 νευρώνες στο επίπεδο εξόδου και να κωδικοποιηθεί για παράδειγμα το ψηφίο 1 με το διάνυσμα εξόδου (1, 0, 0, 0, 0, 0, 0, 0, 0, 0), το ψηφίο 2 με το διάνυσμα εξόδου (0, 1, 0, 0, 0, 0, 0, 0, 0, 0), κοκ. Τέτοιου είδους προσεγγίσεις, η αντιστοίχιση δηλ. Μιας κατηγορίας σε έναν νευρώνα εξόδου, είναι περισσότερο δημοφιλείς όταν ο αριθμός των κατηγοριών είναι γνωστός.

Πρέπει να τονιστεί εδώ ότι η χρήση ενός μόνο νευρώνα εξόδου για αντιμετώπιση προβλημάτων κατηγοριοποίησης στα οποία οι κατηγορίες δεν έχουν κάποια σειρά (unordered categories) πρέπει να αποφεύγεται, καθώς μπορεί να προκύψουν τελείως λανθασμένα αποτελέσματα.

Υποδείξεις σαν τις παραπάνω, δεν υπάρχουν όμως για τα κρυφά επίπεδα. Ο αριθμός των νευρώνων στα κρυφά επίπεδα σχετίζεται με πολύπλοκο τρόπο με παραμέτρους όπως τον αριθμό των νευρώνων στα επίπεδα εισόδου και εξόδου, τον αριθμό των διανυσμάτων εκπαίδευσης και την ύπαρξη ή όχι θορύβου σε αυτά, την πολυπλοκότητα της συνάρτησης ή της κατηγοριοποίησης που καλείται να μάθει το ΤΝΔ, τις συναρτήσεις ενεργοποίησης που χρησιμοποιούνται, τον αλγόριθμο εκπαίδευσης και άλλες.

Ευτυχώς, ανάλογα με το πρόβλημα, υπάρχουν εμπειρικοί κανόνες που βάζουν κάποια όρια στην αρχιτεκτονική του δικτύου που θα χρησιμοποιηθεί. Για παράδειγμα, ο αριθμός των νευρώνων στα κρυφά επίπεδα σε προβλήματα κατηγοριοποίησης είναι καλό να είναι μικρότερος από τον αριθμό των διανυσμάτων εκπαίδευσης ώστε να αποφευχθούν φαινόμενα απομνημόνευσης, η αποκλειστική συσχέτιση δηλαδή ενός κρυφού νευρώνα με ένα διάνυσμα εκπαίδευσης. Γενικά, αποτελεί κοινή παραδοχή ότι απαιτούνται αρκετές δοκιμές και πειραματισμοί μέχρις ότου βρεθεί μία καλή δομή δικτύου για δεδομένο πρόβλημα.

Κανόνας δεν υπάρχει επίσης για την συνδεσμολογία μεταξύ των νευρώνων, αν και η μέθοδος εκπαίδευσης θέτει περιορισμούς σε αυτή. Στην πράξη, αρκετά συνηθισμένη είναι η περίπτωση όπου κάθε νευρώνας συνδέεται με όλους τους νευρώνες του επόμενου επιπέδου. Είναι προφανές ότι ο αριθμός των συνδέσεων μπορεί να είναι αρκετά μεγάλος, με την ακραία περίπτωση του μέγιστου αριθμού συνδέσεων για δίκτυο N νευρώνων να φτάνει τις N(N - 1) συνδέσεις με κατεύθυνση (ή N(N-1)/2 αν αγνοηθεί η κατεύθυνση των συνδέσεων)

Perceptron -> Η πιο απλή μορφή δικτύου πρόσθιας τροφοδότησης.

Δύο άλλοι γνωστοί αλγόριθμοι εκπαίδευσης ΤΝΔ είναι ο κανόνας Δέλτα για δίκτυα χωρίς κρυφά επίπεδα και η ανάστροφη μετάδοση λάθος για ΤΝΔ με κρυφά επίπεδα.

## 2.3 Τα δίκτυα Perceptron

Το πιο απλό νευρωνικό δίκτυο που μπορεί να σχεδιαστεί και να μελετηθεί είναι ασφαλώς ένα δίκτυο που αποτελείται από ένα μόνο νευρώνα. Η λέξη δίκτυο σε αυτή την περίπτωση χρησιμοποιείται καταχρηστικά αφού δεν υπάρχουν περισσότεροι του ενός νευρώνες για να συνδεθούν μεταξύ τους.  
  
Όταν 2 κατηγορίες βρίσκονται αρκετά μακριά η μία από την άλλη και μπορεί να βρεθεί μία ευθεία γραμμή η οποία να τις διαχωρίζει, τότε λέμε ότι οι κλάσεις είναι γραμμικά διαχωρίσιμες. Όταν υπάρχει μία τέτοια ευθεία, τότε υπάρχει και ένα δίκτυα Perceptron το οποίο αντιστοιχεί σ’ αυτή τη γραμμή.

Το ζητούμενο σε ένα νευρωνικό δίκτυο όπως το Perceptron είναι η αυτόματη εκμάθηση των παραμέτρων του συστήματος ώστε να επιτυγχάνεται ο επιθυμητός στόχος, που στην συγκεκριμένη περίπτωση είναι η εύρεση της διαχωριστικής γραμμής. Υποθέτουμε ότι δεν υπάρχει καμία εξωτερική “αυθεντία” η οποία γνωρίζει τις ορθές παραμέτρους για τη λύση του προβλήματος. Ωστόσο το δίκτυο εκπαιδεύεται με επίβλεψη, δηλαδή υπάρχει ένας “δάσκαλος” που μας δίνει την τιμή στόχου για κάθε πρότυπο εκπαίδευσης. Το δίκτυο μαθαίνει προσαρμόζοντας τις παραμέτρους λαμβάνοντας υπ’ όψη του τα επαυξημένα πρότυπα εκπαίδευσης και τους στόχους των προτύπων αυτών χρησιμοποιώντας κάποιον επαναληπτικό αλγόριθμο.

Ο κλασικός κανόνας εκπαίδευσης Perceptron είναι γνωστός και ως κανόνας σταθερής αύξησης (fixed increment rule). Ο κανόνας είναι επαναληπτικός: Τα πρότυπα παρουσιάζονται στο δίκτυο με κυκλική σειρά και όταν τελειώσουν επαναλαμβάνονται από την αρχή. Ένας πλήρης κύκλος χρήσης όλων των προτύπων καλείται εποχή (epoch).

Αποδεικνύεται ότι το μοντέλο Perceptron εφοδιασμένο με τον αλγόριθμο εκπαίδευσης που περιγράψαμε παραπάνω συγκλίνει σε μια λύση η οποία ταξινομεί σωστά όλα τα πρότυπα αρκεί να υπάρχει μια τέτοια λύση, αρκεί, με άλλα λόγια, το πρόβλημα να είναι γραμμικά διαχωρίσιμο. Αν το πρόβλημα δεν είναι γραμμικά διαχωρίσιμο τότε ο αλγόριθμος Perceptron δεν συγκλίνει ποτέ. Αυτό είναι ένα σοβαρό μειονέκτημα το οποίο αποτέλεσε και το βασικό σημείο κριτικής εναντίον του Perceptron. Επιπλέον, πολλά προβλήματα στον πραγματικό κόσμο είναι μη γραμμικά διαχωρίσιμα και επομένως το μοντέλο δεν μπορεί να χρησιμοποιηθεί σε αυτά.

Είναι αυτοπροσαρμοστικός αλγόριθμος.  
  
Η συνάρτηση XOR είναι ίσως η πιο χαρακτηριστική και απλή περίπτωση συνάρτησης που δεν μπορεί να υλοποιηθεί από το δίκτυο Perceptron.

Τo perceptron είναι μια απλή τοπολογία δικτύου πρόσθιας τροφοδότησης χωρίς κρυφά επίπεδα και αποτελεί ιστορικά μια πρώτη προσέγγιση τεχνητών νευρωνικών δικτύων. Προτάθηκε το 1958 από τον Resenblatt ως ένας μηχανισμός που μπορεί να εκπαιδευτεί στην κατηγοριοποίηση προτύπων και σε διάφορες παραλλαγές, πιο απλές ή πιο σύνθετες, εξακολουθεί να υφίσταται ως σήμερα.

Η πιο απλή μορφή Perceptron είναι το στοιχειώδες perceptron (elementary perceptron), το οποίο περιλαμβάνει έναν και μοναδικό τεχνητό νευρώνα, ο οποίος χρησιμοποιεί ως συνάρτηση ενεργοποίησης τη βηματική συνάρτηση. Η μάθηση στο στοιχειώδες perceptron είναι καθοδηγούμενη από το σφάλμα (error driven) και συνίσταται στον υπολογισμό κατάλληλων τιμών βαρών έτσι ώστε δεδομένου ενός δυαδικού διανύσματος εισόδου (που αποτελείται δηλ. από 0 και 1) να παραχθεί η επιθυμητή έξοδος t. Πρόκειται δηλ. Για μία απλή μορφή μάθησης με επίβλεψη. Οι αρχικές τιμές για τα βάρη ορίζονται τυχαία, συνήθως στο διάστημα 0 μέχρι 1, ενώ ο αλγόριθμος μεταβολής των βαρών (αλγόριθμος μάθησης perceptron) έχει ως εξής:



Στον παραπάνω γενικό αλγόριθμο, η ποσότητα d που ονομάζεται ρυθμός μάθησης (learning rate), έχει συνήθως τιμή μεταξύ 0 και 1 και καθορίζει τον ρυθμό μεταβολής των βαρών ο οποίος μάλιστα λαμβάνει χώρα μόνο όταν η υπολογιζόμενη έξοδος y είναι διαφορετική απο την επιθυμητή έξοδο t. Είναι φανερό ότι μόνο βάρη που επιδρούν σε σήμα εισόδου διάφορο του μηδενός θα υποστούν μεταβολή (εκπαίδευση), καθώς διαφορετικά θα είναι x=0 άρα και Dw = 0. Να σημειωθεί ότι η σχέση μεταβολής των βαρών, στη γενική της μορφή, ισχύει και για την περίπτωση που οι τιμές εισόδου/εξόδου δεν είναι δυαδικές (0 ή 1) αλλά διπολικές (-1 ή 1).

Αποδεικνύεται ότι έαν υπάρχει ένα διάνυσμα βαρών που παράγει την επιθυμητή έξοδο για όλα τα διανύσματα εκπαίδευσης, τότε ξεκινώντας από ένα τυχαίο διάνυσμα βαρών και μετά από πεπερασμένο αριθμό βημάτων, ο αλγόριθμος perceptron θα συγκλίνει σε κάποιο διάνυσμα βαρών , όχι απαραίτητα το το οποίο επίσης θα παράγει την επιθυμητή έξοδο για όλα τα διανύσματα εκπαίδευσης. Με άλλα λόγια, ο αλγόριθμος συγκλίνει για κάθε πρόβλημα που μπορεί να αναπαρασταθεί με perceptron. Η χαρακτηριστική ιδιότητα αυτής της κατηγορίας προβλημάτων ονομάζεται γραμμική διαχωρισιμότητα.

Με βάση το στοιχειώδες perceptron, είναι δυνατό να αναπτυχθούν περισσότερο προχωρημένα μοντέλα perceptron που αναγκαστικά θα περιέχουν περισσότερους από ένα νευρώνες. Ο μηχανισμός λειτουργίας τους όμως παραμένει ίδιος με αυτόν του στειχειώδους perceptron.

## 2.4 Τα δίκτυα Adaline

Ο όρος Adaline προέρχεται από τα αρχικά των λέξεων ADaptive LINear Element (Αυτοπροσαρμοζόμενο Γραμμικό Στοιχείο). Τον εισήγαγε ο Windrow για να περιγράψει ένα απλοποιημένο μοντέλο του νευρώνα όπου εκλείπει τελέιως η μη γραμμική συνάρτηση. Η διαφορά σε σχέση με τον κλασικό μη γραμμικό νευρώνα των McCulloch-Pitts είναι ότι τώρα η έξοδος παίρνει συνεχείς τιμές και όχι διακριτές (0/1 ή -1/1). Αυτό σημαίνει ότι και οι στόχοι μπορεί (αλλά δεν είναι και υποχρεωτικό) να παίρνουν τιμές από το σύνολο των πραγματικών αριθμών. Το πρόβλημα είναι πως οι τιμές των στόχων δεν είναι προφανείς αν δεν προέρχονται από ένα διακριτικό σύνολο. Αν επιθυμούμε να χρησιμοποιήσουμε το δίκτυο Adaline για να διαχωρίσουμε δύο κλάσεις, ένα δεύτερο πρόβλημα που πρέπει να αντιμετωπίσουμε είναι να ορίσουμε πότε η έξοδος δείχνει ότι το πρότυπο ανήκει στην πρώτη κατηγορία και πότε στη δεύτερη, καθώς μπορεί να πάρει άπειρες τιμές. Το πρόβλημα λύνεται εύκολα αν για παράδειγμα, πούμε πως οι θετικές τιμές ανήκουν στην πρώτη κλάση, και όλες οι υπόλοιπες στην δεύτερη κλάση.

Στην πράξη ο αλγόριθμος Adaline εξομοιώνεται στον υπολογιστή με πεπερασμένο αριθμό επαναλήψεων, και φυσικά τα πρότυπα εισόδου έχουν επίσης πεπερασμένο πλήθος. Αν το πλήθος των προτύπων είναι μικρό, τότε τα χρησιμοποιούμε με κυκλική επανάληψη έτσι ώστε να δημιουργήσουμε μία ακολουθία με άπειρο μήκος. Όπως και στον αλγόριθμο Perceptron, μία επανάληψη όλων των προτύπων λέγεται εποχή.

Είναι αυτοπροσαρμοστικός αλγόριθμος. Δε χρειάζεται το πρόβλημα να είναι γραμμικά διαχωρίσιμο.

Βασικό μειονέκτημα είναι ότι δεν εγγυάται το διαχωρισμό των κλάσεων όταν το πρόβλημα είναι γραμμικά διαχωρίσιμο. Γενικότερα υπάρχει πρόβλημα στην επιλογή των στόχων καθώς αυτοί είναι πραγματικοί αριθμοί. Στον αλγόριθμο Perceptron τέτοιο πρόβλημα δεν υφίσταται καθώς οι στόχοι είναι δυαδικοί και η επιλογή του στόχου είναι απλή. Εξαρτάται μόνο από την κλάση στην οποία ανήκει το πρότυπο.

## 2.5 Τα δίκτυα MLP

Το δίκτυo MLP έχει απεριόριστες δυνατότητες αναπαράστασης συναρτήσεων και διαχωρισμού κλάσεων.

Προκειμένου να υλοποιήσουμε την πύλη eXclusive OR (XOR), δεν υπάρχει καμία ευθεία γραμμή που να διαχωρίζει τις δύο κλάσεις άρα το δίκτυο Perceptron αδυνατεί να λύσει αυτό το πρόβλημα. Αυτό οδήγησε στην μελέτη του δικτύου Perceptron πολλών στρωμάτων.

Όπως είδαμε οι δυνατότητες αναπαράστασης διαχωριστικών επιφανειών είναι περιορισμένες στο δίκτυο Perceptron καθώς με ένα μόνο νευρώνα το δίκτυο μπορεί να αναπαραστήσει μόνο επίπεδες επιφάνειες. Ο περιορισμός αυτός αίρεται με τη χρήση περισσότερων νευρώνων. Δίκτυα τέτοιου τύπου καλούνται δίκτυα Perceptron πολλών στρωμάτων (Multi-Layer Perceptron - MLP). Το χαρακτηριστικό των δικτύων αυτών είναι ότι οι νευρώνες του οποιουδήποτε στρώματος l τροφοδοτούν αποκλειστικά τους νευρώνες του επόμενου στρώματος (l + 1) και τροφοδοτούνται αποκλειστικά από τους νευρώνες του προηγούμενου στρώματος (l - 1). Για λόγους ευκολίας καλούμε το στρώμα εισόδου καταχρηστικά, μηδενικό στρώμα. Οπότε ο κανόνας γενικεύεται και για το πρώτο στρώμα, δηλαδή το πρώτο στρώμα τροφοδοτείται και αυτό αποκλειστικά από το μηδενικό στρώμα ή από το στρώμα εισόδου.

Τα MLP δίκτυα όπου οι νευρώνες χρησιμοποιούν τη βηματική συνάρτηση 0/1 ή -1/1, μπορούν να υλοποιήσουν συναρτήσεις που δεν μπορεί να υλοποιήσει ένα απλό δίκτυο Perceptron. Ωστόσο, η χρήση της βηματικής συνάρτησης δεν προτιμάται διότι οι περισσότεροι κανόνες εκπαίδευσης βασίζονται σε μεθόδους βελτιστοποίησης, όπως πχ η μέθοδος της κατάβασης δυναμικού, οι οποίες χρησιμοποιούν παραγώγους, ενώ η βηματική συνάρτηση δεν είναι παραγωγίσιμη. Αυτή είναι μία τεχνική δυσκολία η οποία παρ’ όλα αυτά ξεπερνιέται με τη χρήση της σιγμοειδούς συναρτήσεως, η οποία είναι παραγωγίσιμη και πρακτικά μοιάζει πολύ με τη βηματική 0/1.

Για τους παραπάνω λόγους, το ενδιαφέρον των επιστημόνων που μελετούν δίκτυα Perceptron πολλών στρωμάτων σχεδόν μονοπωλείται από δίκτυα των οποίων οι νευρώνες χρησιμοποιούν τη σιγμοειδή συνάρτηση. Μία άλλη συνάρτηση που χρησιμοποιείται εναλλακτικά και μοιάζει με τη βηματική συνάρτηση -1/1 είναι η λεγόμενη υπερβολική εφαπτομένη, η οποία είναι παραγωγίσιμη.

Η χρήση των “μαλακών” συναρτήσεων κατωφλίωσης όπως η σιγμοειδής συνάρτηση ή υπερβολική εφαπτομένη δημιουργεί ομαλές επιφάνειες χωρίς απότομες μεταβολές στην τιμή της εξόδου του δικτύου.

Τα δίκτυα MLP που ενσωματώνουν τη σιγμοειδή συνάρτηση αποδεικνύεται ότι έχουν πολλές δυνατότητες αναπαράστασης συναρτήσεων. Το βασικό θεώρημα ουσιαστικά λέει ότι δίκτυα αυτής της μορφής μπορούν να προσεγγίσουν οποιαδήποτε ομαλή συνάρτηση, όσο κοντά επιθυμούμε. Για το λόγο αυτό τα δίκτυα αυτά καλούνται και “Universal Approximators” (Καθολικοί Προσεγγιστές). Το ενδιαφέρον είναι ότι αρκεί να χρησιμοποιηθούν μόλις δύο στρώματα νευρώνων πέρα από το στρώμα εισόδου.

Το MLP δίκτυο μπορεί να υλοποιήσει οποιαδήποτε συνεχή διαχωριστική επιφάνεια σε n διαστάσεις σε αντίθεση με το απλό δίκτυο Perceptron το οποίο μπορεί να υλοποιήσει μόνο γραμμικές επιφάνειες.

## 2.6 Η μέθοδος Back Propagation

Αλγόριθμος ικανός να εκπαιδεύσει ένα δίκτυο με περισσότερους από έναν νευρώνες.

Η εκπαίδευση ενός δικτύου MLP είναι η διαδικασία ρύθμισης των συναπτικών βαρών του έτσι ώστε να ικανοποιείται κάποιο κριτήριο καταλληλότητας. Άλλωστε αυτός είναι και ο στόχος της εκπαίδευσης σε οποιοδήποτε νευρωνικό δίκτυο όπως για παράδειγμα σε ένα απλό Perceptron. Αυτό που κάνει την εκπαίδευση ενός MLP δικτύου πολύ πιο ενδιαφέρουσα είναι η ιδιότητα του καθολικού προσεγγιστή. Με απλά λόγια, αυτή λέει πως αν έχουμε το κατάλληλο σε μέγεθος δίκτυο μπορούμε να το εκπαιδεύσουμε να μάθει οποιαδήποτε συνάρτηση εμείς επιθυμούμε. Θυμίζουμε σε αντιδιαστολή, ότι το απλό Perceptron μπορεί να υλοποιήσει μόνο γραμμικές διαχωριστικές επιφάνειες. Κυριότερος εκπρόσωπος των αλγορίθμων εκπαίδευσης MLP είναι ο Back Propagation.

O αλγόριθμος Back-Propagation προτάθηκε από τον Paul Werbos στη δεκαετία του 1970 στα πλαίσια της ανάλυσης μοντέλων οικονομικής και πολιτικής πρόβλεψης. Τότε, αυτά τα μοντέλα δεν είχαν σχεδιαστεί για να έχουν σχέση ή να θυμίζουν νευρωνικά δίκτυα. Παρόμοια μοντέλα είχαν προταθεί ανεξάρτητα και από άλλους ερευνητές όπως οι Bryson & Ho. Τη δεκαετία του 80, έγινε αντιληπτό ότι η μέθοδος μπορούσε να μεταφερθεί αυτούσια στην εκπαίδευση νευρωνικών δικτύων MLP και έκτοτε έγινε η πιο γνωστή και η πιο διαδεδομένη μέθοδος.

Βασικό χαρακτηριστικό της μεθόδου αυτής είναι η ύπαρξη στόχων, όπως ακριβώς και στο απλό δίκτυο Perceptron. Συνεπώς το μοντέλο ανήκει στην κατηγορία των δικτύων που εκπαιδεύονται με επίβλεψη.

Ο αλγόριθμος αυτός έχει στην πράξη διάφορα μειονεκτήματα, το βασικότερο των οποίων είναι η αργή σύγκλιση. Το πρόβλημα των τοπικών ελαχίστων είναι επίσης κρίσιμο αλλά πρακτικά η αντιμετώπιση του είναι ιδιαίτερα δύσκολη. Διάφορες μέθοδοι έχουν αναπτυχθεί με στόχο την επιτάχυνση της σύγκλισης όπως:

1. Η χρήση της ορμής (momentum)
2. Η αναζήτηση σε μία ευθεία γραμμή (line search)  
   3. Η συζυγής κατάβαση δυναμικού (conjugate gradient)

Και οι διάφορες παραλλαγές τις μεθόδου Newton.

## 2.7 Η μέθοδος Gradient Descent

Η κατάβαση δυναμικού (gradient descent) είναι μια κλασική μέθοδος εύρεσης της ελάχιστης τιμής μίας συνάρτησης κόστους n μεταβλητών. Αν επιθυμούμε την εύρεση του μέγιστου σημείου της συνάρτησης τότε η ίδια μέθοδος μπορεί να χρησιμοποιηθεί πάλι με μόνη διαφορά το πρόσημο της κατεύθυνσης αναζήτησης του βέλτιστου σημείου. Στη περίπτωση αυτή, η μέθοδος καλείται ανάβαση δυναμικού (gradient ascent).

# Γενετικοί αλγόριθμοι

Οι γενετικοί αλγόριθμοι είναι αλγόριθμοι αναζήτησης που βασίζονται στη μηχανική της φυσικής επιλογής και της φυσικής γενετικής. Συνδυάζουν την επιβίωση του καταλληλότερου μεταξύ των δομών συμβολοσειρών με μια δομημένη αλλά τυχαιοποιημένη ανταλλαγή πληροφοριών για να σχηματίσουν έναν αλγόριθμο αναζήτησης με κάποιο από τα καινοτόμα ίχνη της ανθρώπινης αναζήτησης. κάθε γενιά, ένα νέο σύνολο τεχνητών πλασμάτων (χορδές) δημιουργείται χρησιμοποιώντας κομμάτια από τα πιο δυνατά από τα παλιά· ένα περιστασιακό νέο μέρος δοκιμάζεται για καλό μέτρο. Αν και είναι τυχαιοποιημένοι, οι γενετικοί αλγόριθμοι δεν είναι απλός τυχαίος περίπατος. Εκμεταλλεύονται αποτελεσματικά ιστορικές πληροφορίες για να κάνουν εικασίες για νέα σημεία αναζήτησης με αναμενόμενη βελτιωμένη απόδοση. Οι γενετικοί αλγόριθμοι αναπτύχθηκαν από τον John Holland, τους συναδέλφους του και τους μαθητές του στο Πανεπιστήμιο του Michigan. Οι στόχοι της έρευνάς τους ήταν δύο: (1) η αφαίρεση και η αυστηρή εξήγηση των προσαρμοστικών διαδικασιών των φυσικών συστημάτων και (2) η σχεδιάστε λογισμικό τεχνητών συστημάτων που διατηρεί τους σημαντικούς μηχανισμούς των φυσικών συστημάτων. Αυτή η προσέγγιση έχει οδηγήσει σε σημαντικές ανακαλύψεις τόσο στην επιστήμη των φυσικών όσο και των τεχνητών συστημάτων.

Το κεντρικό θέμα της έρευνας για τους γενετικούς αλγόριθμους ήταν η ευρωστία, η ισορροπία μεταξύ της αποτελεσματικότητας και της αποτελεσματικότητας που απαιτείται για την επιβίωση σε πολλές διαφορετικές. Οι γενετικοί αλγόριθμοι είναι θεωρητικά και εμπειρικά αποδεδειγμένα

παρέχουν εύρωστη αναζήτηση σε σύνθετους χώρους.Η πρωτογενής μονογραφία για το θέμα

είναι το Holland's(1975)Adaptation in Natural and Artificial Systems.Πολλές εργασίες

και οι διατριβές καθιερώνουν την εγκυρότητα της τεχνικής σε εφαρμογές βελτιστοποίησης συναρτήσεων και ελέγχου. Έχοντας καθιερωθεί ως έγκυρη προσέγγιση σε προβλήματα που απαιτούν αποτελεσματική και αποτελεσματική αναζήτηση, οι γενετικοί αλγόριθμοι βρίσκουν τώρα πιο διαδεδομένη εφαρμογή σε επιχειρηματικούς, επιστημονικούς και μηχανικούς κύκλους. Οι αλγόριθμοι πίσω από τον αυξανόμενο αριθμό εφαρμογών είναι ξεκάθαροι. Αυτοί οι αλγόριθμοι είναι υπολογιστικά απλοί αλλά ισχυροί στην αναζήτησή τους για βελτίωση. Επιπλέον, δεν περιορίζονται θεμελιωδώς από περιοριστικές υποθέσεις σχετικά με τον χώρο αναζήτησης (υποθέσεις σχετικά με τη συνέχεια, την ύπαρξη παραγώγων, τροπικότητα, και άλλα θέματα). Θα διερευνήσουμε τους λόγους πίσω από αυτές τις ελκυστικές ιδιότητες, αλλά πριν από αυτό, πρέπει να διερευνήσουμε την ευρωστία των πιο ευρέως αποδεκτών διαδικασιών αναζήτησης. (Goldberg, D. E., 1989) Οι γενετικοί αλγόριθμοι (ΓΑ) (genetic algorithms (GA)) αποτελούν τους δημοφιλέστερους από τους εξελικτικούς αλγορίθμους και βασίζονται σε αρχές της εξέλιξης των ειδών που εισήγαγε ο Δαρβίνος με το βιβλίο του «Η Καταγωγή των Ειδών» (Darwin, 2014). Οι ΓΑ προτάθηκαν τη δεκαετία του 1960 από τον Holland (1992) και τους συνεργάτες του. Παρά το γεγονός ότι η αρχική ορολογία που εισήγαγε ο Holland ήταν γενετικό σχέδιο (genetic plan), ο όρος γενετικός αλγόριθμος υιοθετήθηκε αργότερα για να υπογραμμίσει τον κεντρικό ρόλο του υπολογισμού στον καθορισμό και την υλοποίηση των σχεδίων. (Καμπουρλάζος &Παπακώστας, 2015)

## 3.1 Μέθοδοι κωδικοποίησης

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Πίνακας 3.2 [Τίτλος πίνακα]

## 3.2 Γενετικοί τελεστές

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Πίνακας 3.2 [Τίτλος πίνακα]

## 3.3 Παράλληλοι γενετικοί αλγόριθμοι

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Πίνακας 3.2 [Τίτλος πίνακα]

## 3.4 Εφαρμογές γενετικών αλγορίθμων

Όπως και τα ΤΝΔ, έτσι και οι γενετικοί αλγόριθμοι αποτελούν έναν εύκολο τρόπο επίλυσης προβλημάτων με μεγάλη δυνατότητα προσαρμογής. Σε αυτό το κεφάλαιο θα αναλύσουμε κάποιες από τις πιο αντιπροσωπευτικές εφαρμογές των γενετικών αλγορίθμων. Μία από αυτές είναι η εύρεση μέγιστης τιμής σε αριθμητικές συναρτήσεις, κάτι που δεν είναι εύκολη υπόθεση για συναρτήσεις πολλών μεταβλητών καθώς εμφανίζουν ασυνέχειες, θόρυβο και άλλα. Μία ακόμη εφαρμογή είναι η επεξεργασία εικόνων, το αποτέλεσμα αυτής μπορεί να αποτελέσει τη βάση για τη μηχανική μάθηση. Μία άλλη εφαρμογή είναι η συνδυαστική βελτιστοποίηση με πιο γνωστό πρόβλημα σε αυτή την κατηγορία, αυτό του πλανόδιου πωλητή κατά το οποίο οι γενετικοί αλγόριθμοι μπορούν να δώσουν αρκετές λύσεις κοντά στη βέλτιστη. Οι γενετικοί αλγόριθμοι μπορούν να χρησιμοποιηθούν στη σχεδίαση κατασκευών και εξαρτημάτων, όπως γέφυρες ή μηχανολογικά εξαρτήματα όπου ζητούμενο μπορεί να είναι η εύρεση μίας λύσης όσο και η βελτιστοποίηση της. Οι αλγόριθμοι μπορούν να δοκιμάσουν συνδυασμούς και ιδέες που ο ανθρώπινος νους δε θα δοκίμαζε ποτέ, δίνοντας ενίοτε αρκετά πρωτότυπα αποτελέσματα. Οι γενετικοί αλγόριθμοι μπορούν να χρησιμοποιηθούν για την προσέγγιση συναρτήσεων με πιο γνωστή εφαρμογή αυτή των συστημάτων ταξινόμησης. Ωστόσο οι γ.α. έχουν χρησιμοποιηθεί και σε παιχνίδια, επίλυση λαβυρίνθων, καθώς και για πολιτικές και οικονομικές αναλύσεις. (Βλαχάβας et al., 2006)

# Μέθοδος - αποτελέσματα

## 4.1 Τα dataset που χρησιμοποιήθηκαν

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Πίνακας 3.1 [Τίτλος πίνακα]

## 4.2 Η προτεινόμενη μέθοδος

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Πίνακας 3.2 [Τίτλος πίνακα]

## 4.3 Πειραματικά αποτελέσματα

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |

Πίνακας 3.2 [Τίτλος πίνακα]

# Συμπεράσματα

Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας. Κυρίως κείμενο της εργασίας.

# ΠΑΡΑΡΤΗΜΑ

*[Υλικό που είναι ογκώδες, παραδείγματος χάριν αποτελέσματα ερωτηματολογίων, σχήματα, πίνακες κ.ά. που εμποδίζουν τον αναγνώστη να κατανοήσει τη συνέχεια του κειμένου, μπορεί να τοποθετηθεί σε παραρτήματα. Η αρίθμηση των παραρτημάτων γίνεται με κεφαλαία ελληνικά γράμμα Α, Β, Γ,… ενώ σχήματα, σχέδια, πίνακες κ.λπ., που περιλαμβάνονται πρέπει να ονομάζονται Σχήμα Α1, Πίνακας Β2, κ.λπ.]*

***Κώδικας που χρησιμοποιήθηκε για την δημιουργία γραφικών παραστάσεων***

Γραμμική συνάρτηση:

x = -10:0.1:10; % Ορισμός των τιμών του x από -10 έως 10 με βήμα 0.1

y = x; % Υπολογισμός των αντίστοιχων τιμών του y

plot(x, y); % Σχεδίαση της γραφικής παράστασης

xlabel('x'); % Ετικέτα x

ylabel('f(x)'); % Ετικέτα y

title('f(x) = x'); % Τίτλος του γραφήματος

grid on; % Ενεργοποίηση του grid

*Βηματική συνάρτηση 0/1:*

*x = -10:0.1:10; % Ορισμός των τιμών του x από -10 έως 10 με βήμα 0.1*

*y = x >= 0; % Υπολογισμός των αντίστοιχων τιμών του y (1 για x >= 0, 0 για x < 0)*

*plot(x, y); % Σχεδίαση της γραφικής παράστασης*

*xlabel('x'); % Ετικέτα x*

*ylabel('f(x)'); % Ετικέτα y*

*title('Step Function'); % Τίτλος του γραφήματος*

*grid on; % Ενεργοποίηση του grid*

*Βηματική συνάρτηση -1/1*

*x = -10:0.1:10; % Ορισμός των τιμών του x από -10 έως 10 με βήμα 0.1*

*y = ones(size(x)); % Δημιουργία ενός πίνακα με μέγεθος ίδιο με τον πίνακα x και τιμές 1*

*y(x < 0) = -1; % Για τις τιμές του x που είναι μικρότερες από το 0, ορίζουμε τιμή -1*

*plot(x, y); % Σχεδίαση της γραφικής παράστασης*

*xlabel('x'); % Ετικέτα x*

*ylabel('f(x)'); % Ετικέτα y*

*title('Step Function (-1/1)'); % Τίτλος του γραφήματος*

*grid on; % Ενεργοποίηση του grid*

*Σιγμοειδής συνάρτηση:*

*x = -10:0.1:10; % Ορισμός των τιμών του x από -10 έως 10 με βήμα 0.1*

*y = 1 ./ (1 + exp(-x)); % Υπολογισμός των αντίστοιχων τιμών του y σύμφωνα με τη σιγμοειδή συνάρτηση*

*plot(x, y); % Σχεδίαση της γραφικής παράστασης*

*xlabel('x'); % Ετικέτα x*

*ylabel('f(x)'); % Ετικέτα y*

*title('Sigmoid Function'); % Τίτλος του γραφήματος*

*grid on; % Ενεργοποίηση του grid*

*Συνάρτηση υπερβολικής εφαπτομένης:*

*x = -10:0.1:10; % Ορισμός των τιμών του x από -10 έως 10 με βήμα 0.1*

*y = tanh(x); % Υπολογισμός των αντίστοιχων τιμών της υπερβολικής εφαπτομένης*

*plot(x, y); % Σχεδίαση της γραφικής παράστασης*

*xlabel('x'); % Ετικέτα x*

*ylabel('tanh(x)'); % Ετικέτα y*

*title('Hyperbolic Tangent Function'); % Τίτλος του γραφήματος*

*grid on; % Ενεργοποίηση του grid*

# ΒΙΒΛΙΟΓΡΑΦΙΑ

# ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Βλαχάβας, Ι., Κεφάλας, Π., Βασιλειάδης, Ν., Κόκκορας, Φ. και Σακελλαρίου, Η. (2006) *Τεχνητή Νοημοσύνη*. 3η έκδοση. Θεσσαλονίκη: Εκδόσεις Πανεπιστημίου Μακεδονίας.

Διαμαντάρας, Κ.(2007) *Τεχνητά Νευρωνικά Δίκτυα*. Αθήνα: Εκδόσεις Κλειδάριθμος.

Θεοδωρίδης, Σ. και Κουτρούμπας, Κ. (2012) *Αναγνώριση προτύπων*. Κύπρος: Εκδόσεις Πασχαλίδη.

Καμπουρλάζος, Β. Γ., Παπακώστας, Γ. Α. (2015) *Εισαγωγή στην υπολογιστική νοημοσύνη*. Αθήνα: Κάλλιππος.

Τσούλος, Ι. (χ.χ.) *Τεχνητά Νευρωνικά Δίκτυα* Διαθέσιμο σε: [URL](https://www.dit.uoi.gr/e-class/modules/document/file.php/249/%CE%94%CE%99%CE%91%CE%9B%CE%95%CE%9E%CE%95%CE%99%CE%A3/lecture1.pdf) (Ημερομηνία πρόσβασης: 08 Ιουνίου 2024).

Haykin, S. (2010). *Νευρωνικά δίκτυα και μηχανική μάθηση.* 3η έκδοση. Αθήνα: Εκδόσεις Παπασωτηρίου.

# ΞΕΝΟΓΛΩΣΣΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Anderson, D. and McNeill, G., (1992). *Artificial neural networks technology*: A DACS State-of-the-Art Report. New York: Kaman Sciences Corporation.

Chiarandini, M. (χ.χ.) *Machine Learning: Linear Regression and Neural Networks.* Διαθέσιμο σε: [URL](https://www.imada.sdu.dk/u/rolf/Edu/DM534/E16/DM534-marco.pdf) (Ημερομηνία πρόσβασης: 08 Ιουνίου 2024).

Goldberg, D. E., (1989) *Genetic Algorithms in Search, Optimization & Machine Learning*. Boston, MA: Addison Wesley Longman, Inc.

Suzuki, K. (ed.) (2011) *Artificial Neural Networks - Methodological advances and biomedical applications.* Croatia: Published by InTech

Russell, S. και Norvig, P. (2010) *Artificial Intelligence: A Modern Approach*. New Jersey: Pearson Educaton, Inc.

Κίτρου, Κ. (χ.χ.) *Τι είναι η μηχανική μάθηση; (Machine Learning)*. Διαθέσιμο σε: [URL](https://www.iekdelta360.gr/ti-einai-i-michaniki-mathisi-machine-learning) (Ημερομηνία πρόσβασης: 08 Ιουνίου 2024).

[Οπισθόφυλλο. Κενή σελίδα]

1. (Διαμαντάρας, 2007) & (Βλαχάβας & Κεφάλας & Βασιλειάδης & Κόκκορας & Σακελλαρίου, 2006) [↑](#footnote-ref-0)
2. (Haykin, 2010) & (Θεοδωρίδης & Κουτρούμπας, 2012) & (Κίτρου, χ.χ.) [↑](#footnote-ref-1)