

**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, 2024

**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.

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](#_Toc3489)

[ΠΕΡΙΛΗΨΗ 10](#_Toc1042)

[ABSTRACT 11](#_Toc30147)

[ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ 12](#_Toc9365)

[ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ 14](#_Toc17315)

[ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ 15](#_Toc32349)

[ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ 16](#_Toc30512)

[ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ 17](#_Toc32671)

[1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα 18](#_Toc972)

[1.1 Ιστορική Αναδρομή 18](#_Toc31818)

[1.2 Μετάβαση από τους βιολογικούς στους τεχνητούς νευρώνες 20](#_Toc11117)

[1.3 Συναρτήσεις ενεργοποίησης 23](#_Toc7452)

[1.4 Τεχνητά Νευρωνικά Δίκτυα: Μία σφαιρική εικόνα 26](#_Toc12281)

[1.5 Μάθηση με επίβλεψη 28](#_Toc20502)

[1.6 Εφαρμογές νευρωνικών δικτύων 30](#_Toc29779)

[2. Δομή και Λειτουργία των Τεχνητών Νευρωνικών Δικτύων 32](#_Toc3536)

[2.1 Τα δίκτυα Perceptron 32](#_Toc13037)

[2.2 Τα δίκτυα Adaline 34](#_Toc8773)

[2.3 Τα δίκτυα MLP 35](#_Toc24036)

[2.4 Η μέθοδος Back Propagation 37](#_Toc23809)

[2.5 Η μέθοδος Gradient Descent 38](#_Toc7401)

[3. Γενετικοί αλγόριθμοι 39](#_Toc27622)

[3.1 Ιστορική Αναδρομή 39](#_Toc16260)

[3.2 Μέθοδοι κωδικοποίησης 40](#_Toc31961)

[3.3 Γενετικοί τελεστές 41](#_Toc28284)

[3.4 Παράλληλοι γενετικοί αλγόριθμοι 42](#_Toc26136)

[3.5 Πλεονεκτήματα γενετικών αλγορίθμων 43](#_Toc32102)

[3.6 Εφαρμογές γενετικών αλγορίθμων 45](#_Toc21004)

[4. Μέθοδος - αποτελέσματα 46](#_Toc1437)

[4.1 Τα dataset που χρησιμοποιήθηκαν 46](#_Toc4813)

[4.2 Η προτεινόμενη μέθοδος 47](#_Toc32324)

[4.3 Πειραματικά αποτελέσματα 48](#_Toc8891)

[5. Συμπεράσματα 49](#_Toc24823)

[ΠΑΡΑΡΤΗΜΑ 51](#_Toc17102)

[ΒΙΒΛΙΟΓΡΑΦΙΑ 56](#_Toc10579)

# 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

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

# ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ

**ΠΡΟΤΥΠΑ:**

**ΜΗ ΓΡΑΜΜΙΚΑ ΔΙΑΧΩΡΙΣΙΜΑ ΠΡΟΤΥΠΑ:** Πρότυπα που δεν μπορούν να διαχωριστούν με μία ευθεία γραμμή κατά την αναπαράσταση τους.

**ΕΠΟΧΗ:** Ένας πλήρης κύκλος χρήσης όλων των προτύπων.

# 1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα

## Ιστορική Αναδρομή

Τη δεκαετία του 1940 σημειώθηκε σημαντική πρόοδος στη μελέτη των βιολογικών νευρωνικών δικτύων και στη μαθηματική μοντελοποίηση τους. Η περίοδος των ΤΝΔ ξεκίνησε το 1943 με την δημοσίευση των Αμερικανών επιστημόνων McCulloch και Pitts, που αφορούσε την ανάπτυξη ενός μοντέλου νευρωνικού δικτύου που φέρει το όνομα τους. Το μοντέλο αυτό συνέδεσε τις επιστήμες της Νευροφυσιολογίας και της Μαθηματικής Ανάλυσης, μιμούμενο τη λειτουργία ενός βιολογικού νευρώνα. Αξιοσημείωτο είναι το γεγονός πως ο von Neumann χρησιμοποίησε τα ιδεατά στοιχεία αυτού του νευρώνα για την κατασκευή του EDVAC, του πρώτου γενικού σκοπού υπολογιστή το 1949.

Εννέα χρόνια αργότερα, το 1958, ο ψυχολόγος Rosenblatt δημοσίευσε το δίκτυο Perceptron, μία καινοτόμα μέθοδο μάθησης με επίβλεψη. Το δίκτυο αυτό, αποτελεί ένα απλό μοντέλο δύο επιπέδων και είναι μία ελαφρώς τροποποιημένη έκδοση του νευρώνα των McCulloch και Pitts. Προτάθηκε ως ένας μηχανισμός που μπορεί να εκπαιδευτεί στην κατηγοριοποίηση προτύπων και εξακολουθεί να υφίσταται μέχρι και σήμερα. Η πρώτη εμπεριστατωμένη κριτική του Perceptron παρουσιάστηκε από τους Minsky και Selfridge κατά την οποία έγινε επισήμανση πως αυτό το δίκτυο δεν θα μπορούσε να υποστηρίξει γενικεύσεις . Έτσι, το 1969 στο περίφημο βιβλίο “Perceptrons” των Minsky και Papert, αποδείχθηκε με μαθηματικό τρόπο, πως ΤΝΔ ενός επιπέδου όπως είναι ο Perceptron, δεν μπορούν να λύσουν προβλήματα που αφορούν μη γραμμικά διαχωρίσιμα πρότυπα. Την ίδια περίοδο, οι Widrow και Hoff, αναπτύσσουν το μοντέλο Adaline, μία τροποποίηση του δικτύου Perceptron και αποτέλεσε ένα από τα πρώτα μοντέλα που εφαρμόστηκαν με επιτυχία σε πρακτικά προβλήματα. Ειδικότερα, χρησιμοποιήθηκε ως φίλτρο για την εξάλειψη της ηχώ σε τηλεφωνικές γραμμές.

Τη δεκαετία του 1970, η έρευνα πάνω στα ΤΝΔ παρουσίασε ύφεση. Όμως, κατά τη δεκαετία του 1980, σημειώθηκε έκρηξη δραστηριότητας με την εισαγωγή του MLP, ο οποίος πρόσφερε απεριόριστες δυνατότητες αναπαράστασης συναρτήσεων και διαχωρισμού κλάσεων. Με αυτό το δίκτυο φαίνεται πως ξεπεράστηκαν και οι υπολογιστικοί περιορισμοί που παρουσιάστηκαν για το δίκτυο Perceptron. Το 1986, οι Rumelhart, Hinton και Williams παρουσίασαν για πρώτη φορά τον αλγόριθμο με οπισθοχώρηση (Back Propagation), ο οποίος έγινε πολύ δημοφιλής και έδωσε νέα ώθηση στις εφαρμογές νευρωνικών δικτύων. Αλγόριθμος, ικανός για πρώτη φορά να εκπαιδεύσει ένα δίκτυο με περισσότερους από έναν νευρώνες. Η ανάπτυξη αυτού του αλγορίθμου αποτέλεσε ορόσημο στην εξέλιξη του κλάδου των νευρωνικών δικτύων.

Τα νευρωνικά δίκτυα σίγουρα έχουν διανύσει πολύ δρόμο από την εποχή των McCulloch και Pitts, και έχουν εγκαθιδρυθεί ως διεπιστημονικό αντικείμενο με βαθιές ρίζες στην ψυχολογία, στα μαθηματικά, στις φυσικές επιστήμες και στη μηχανική. [[1]](#footnote-0)

## 1.2 Μετάβαση από τους βιολογικούς στους τεχνητούς νευρώνες

Ο ανθρώπινος εγκέφαλος είναι το πιο πολύπλοκο όργανο του ανθρώπινου σώματος. Επιτελεί τις ανώτερες λειτουργίες, δέχεται αισθητικά ερεθίσματα και εκπέμπει κινητικές διεγέρσεις.[[2]](#footnote-1) O Ramón y Cajal ήταν ο πρώτος που βοήθησε προκειμένου να γίνει πιο εύκολη η κατανόηση του εγκεφάλου, εισάγοντας την ιδέα των νευρώνων ως δομικά συστατικά του. Ο νευρώνας είναι ένα μεγάλο σε μέγεθος κύτταρο, και παρά την ύπαρξη αρκετών παραλλαγών του, ανατομικά αποτελείται από τέσσερα βασικά συστατικά. Τους δενδρίτες, το σώμα - που αποτελεί και τον πυρήνα του νευρώνα, 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. Σε κάθε δενδρίτη υπάρχει ένα απειροελάχιστο κενό που ονομάζεται σύναψη. Οι συνάψεις αυτές, με τη σειρά τους - συχνά αναφέρονται και ως νευρικές απολήξεις, είναι στοιχειώδεις δομικές και λειτουργικές μονάδες που παίζουν διαμεσολαβητικό ρόλο κατά τις αλληλεπιδράσεις μεταξύ των νευρώνων. Στόχος τους είναι η μετάδοση της ηλεκτρικής δραστηριότητας από τον άξονα - αποστολέα στους δενδρίτες - παραλήπτες, δημιουργώντας έτσι ένα νευρωνικό δίκτυο. Κάθε σύναψη έχει ένα ειδικό βάρος που καθορίζει τη δύναμη της και ονομάζεται συναπτικό βάρος. Κάθε νευρώνας αξιολογεί τα εισερχόμενα σήματα χρησιμοποιώντας τα αντίστοιχα συναπτικά βάρη προκειμένου να υπολογίσει το τελικό φορτίο. Όσο πιο ισχυρή είναι η σύνδεση μεταξύ τους, 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. Αν όμως το άθροισμα του φορτίου δεν ξεπερνάει το κατώφλι, τότε ο νευρώνας εκπέμπει σε τυχαία χρονικά διαστήματα, πολύ αραιά παλμούς και λέμε ότι ο νευρώνας είναι αδρανής και έτσι θεωρούμε τη σύναψη ως ανασταλτική. (Βλέπε Εικόνα 1.1 για τη δομή του βιολογικού νευρώνα)[[3]](#footnote-2)

Ο εγκέφαλος έχει την ικανότητα να λαμβάνει πολύπλοκες αποφάσεις με εκπληκτική ταχύτητα. Παρά το γεγονός ότι οι νευρώνες είναι πιο αργοί από τις λογικές πύλες που υλοποιούνται με τεχνολογία πυριτίου, αυτό επιτυγχάνεται καθώς η υπολογιστική ικανότητα του αλλά 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. Αξίζει να σημειωθεί πως οι τεχνητοί νευρώνες που χρησιμοποιούμε για την κατασκευή των ΤΝΔ είναι αρκετά πρωτόγονοι σε σύγκριση με τους νευρώνες του ανθρώπινου εγκεφάλου. Ο τεχνητός νευρώνας λοιπόν, 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. Το σώμα του τεχνητού νευρώνα χωρίζεται σε δύο τμήματα. Το πρώτο αποτελείται από τον αθροιστή (adder), ο οποίος προσθέτει τα επηρεασμένα από τα βάρη σήματα εισόδου. Το δεύτερο τμήμα αποτελείται από τη συνάρτηση ενεργοποίησης (activation function), η οποία λειτουργεί ως φίλτρο και στόχος της είναι η διαμόρφωση της τελικής τιμής του σήματος εξόδου. Συχνά αναφέρεται και ως συνάρτηση περιορισμού (squashing function), καθώς περιορίζει το επιτρεπτό εύρος πλάτους του σήματος εξόδου σε κάποια πεπερασμένη τιμή. 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. Συνήθως 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 βάρη, 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. (Βλέπε Εικόνα 1.2 για τη δομή του τεχνητού νευρώνα)[[4]](#footnote-3)

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



(Εικόνα 1.2) Τεχνητός νευρώνας (Suzuki, 2011)

## 1.3 Συναρτήσεις ενεργοποίησης

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



Γραμμική συνάρτηση:[[5]](#footnote-4)

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

(Εικόνα 1.3) Γραφική παράσταση της γραμμικής συνάρτησης

Βηματική συνάρτηση:[[6]](#footnote-5)



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

Αλλά και με την μορφή

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

Σιγμοειδής συνάρτηση:[[7]](#footnote-6)



(Εικόνα 1.6) Γραφική παράσταση της σιγμοειδής συνάρτησης



Υπερβολική εφαπτομένη:

(Εικόνα 1.7) Γραφική παράσταση της συνάρτησης υπερβολικής εφαπτομένης

## 1.4 Τεχνητά Νευρωνικά Δίκτυα: Μία σφαιρική εικόνα

Τα ΤΝΔ αποτελούν μία ιδιαίτερη προσέγγιση στη δημιουργία συστημάτων με νοημοσύνη επειδή βασίζονται σε βιολογικά πρότυπα καθώς χρησιμοποιούν δομές και διαδικασίες που μιμούνται τις αντίστοιχες του ανθρώπινου εγκεφάλου. Από τεχνικής άποψης είναι συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος τεχνητών νευρώνων, οργανωμένων σε δομές παρόμοιες με τις βιολογικές. Μπορούμε να αναφέρουμε δύο (2) στοιχεία στα οποία μοιάζει ένα ΤΝΔ με τον ανθρώπινο εγκέφαλο. Το πρώτο από αυτά είναι πως το δίκτυο προσλαμβάνει τη γνώση από το περιβάλλον του, μέσω μίας διαδικασίας μάθησης και το δεύτερο είναι πως η ισχύς των συνδέσεων μεταξύ των νευρώνων, χρησιμοποιείται για την αποθήκευση της γνώσης που αποκτιέται. Έτσι, σύμφωνα με τον Βλαχάβα, μπορεί να διατυπωθεί ο παρακάτω ορισμός:

*Ένα νευρωνικό δίκτυο 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.*

Χάριν συντομίας, τα δίκτυα αυτά σύμφωνα με τον Haykin, αποκαλούνται και νευρωνικά δίκτυα. Αυτές οι μονάδες επεξεργασίας, δηλαδή οι νευρώνες, συνήθως είναι οργανωμένοι σε μία σειρά από στρώματα ή αλλιώς επίπεδα (layers). Το πρώτο από αυτά ονομάζεται επίπεδο εισόδου (input layer) και χρησιμοποιείται για την εισαγωγή των δεδομένων. Η χρήση του όρου νευρώνες σε αυτό το επίπεδο είναι καταχρηστική καθώς δεν γίνεται κάποια επεξεργασία σε αυτούς. Στη συνέχεια, υπάρχουν προαιρετικά ένα ή περισσότερα κρυφά επίπεδα (hidden layer) ενώ στο τέλος υπάρχει το επίπεδο εξόδου (output layer).

Οι νευρώνες των δικτύων αυτών, μπορούν να είναι πλήρως ή μερικώς συνδεδεμένοι. Όταν συνδέονται με όλους τους υπόλοιπους νευρώνες αναφερόμαστε σε πλήρως συνδεδεμένους νευρώνες, ενώ σε διαφορετική περίπτωση αναφερόμαστε σε μερικώς συνδεδεμένους νευρώνες. Μίας συνήθης περίπτωση των μερικώς συνδεδεμένων νευρώνων, είναι τα δίκτυα πρόσθιας τροφοδότησης (feedforward) κατά τα οποία, οι νευρώνες ενός επιπέδου είναι πλήρως διασυνδεδεμένοι με αυτούς του επόμενου επιπέδου, αλλά όχι με αυτούς του προηγούμενου και χρησιμοποιούνται στην πλειοψηφία των εφαρμογών των ΤΝΔ.

Υπάρχουν τέσσερις (4) ιδιότητες που είναι άρρηκτα συνδεδεμένες με τα ΤΝΔ. Η πρώτη από αυτές είναι η ικανότητα τους να μαθαίνουν μέσω παραδειγμάτων (learn by example). Η δεύτερη είναι η δυνατότητα θεώρησης τους ως κατανεμημένη μνήμη (distributed memory) και ως μνήμη συσχέτισης (associative memory). Αυτό πηγάζει από το γεγονός ότι η κωδικοποίηση που δημιουργούν είναι κατανεμημένη σε όλα τα βάρη της συνδεσμολογίας τους. Μία μνήμη συσχέτισης δεν αποθηκεύει πληροφορία με τον παραδοσιακό τρόπο αλλά μέσω κατάλληλων συσχετίσεων που δημιουργεί από τα δεδομένα εκπαίδευσης. Έτσι, η ανάκληση της πληροφορίας γίνεται βάσει του περιεχομένου και όχι της διεύθυνσης, όπως συμβαίνει στον ανθρώπινο εγκέφαλο. H τρίτη ιδιότητα είναι η μεγάλη τους ανοχή σε σφάλματα (fault-tolerant) καθώς ορισμένα είδη ΤΝΔ είναι σε θέση να παράγουν τη σωστή έξοδο ακόμη και αν τα δεδομένα εισόδου είναι λίγο διαφορετικά. Μία τέτοια κατάσταση μπορεί να προκληθεί είτε λόγω θορύβου είτε επειδή τα δεδομένα είναι ελλειπή. Αυτό ισχύει ακόμη και για δομικά σφάλματα, κάτι το οποίο σημαίνει πως η κακή λειτουργία ή ακόμη και η καταστροφή ενός νευρώνα ή κάποιων συνδέσεων δεν είναι ικανή να διαταράξει σημαντικά τη λειτουργία του καθώς η πληροφορία που εσωκλείουν δεν είναι εντοπισμένη σε συγκεκριμένο σημείο αλλά είναι διάχυτη σε όλο το δίκτυο. Η τελευταία ιδιότητα είναι η ικανότητα τους για αναγνώριση προτύπων (pattern recognition). Από τη στιγμή που ένα ΤΝΔ εκπαιδευτεί στο να αναγνωρίζει συνθήκες και καταστάσεις, τότε προκειμένου να προσδιοριστεί μία συγκεκριμένη κατάσταση, απαιτείται ένας μόνο κύκλος λειτουργίας του. 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, ακόμη και σε πεδία μάχης.[[8]](#footnote-7)

## 1.5 Μάθηση με επίβλεψη

Τα ΤΝΔ βασίζονται στη λειτουργία του ανθρώπινου εγκεφάλου, μιμούμενα τόσο τη λειτουργία των βιολογικών νευρώνων όσο και τη δομή των βιολογικών νευρωνικών δικτύων. Συνεπώς, η πρόκληση που αντιμετωπίζουν είναι η εύρεση αλγορίθμων κατάλληλων με σκοπό την εκπαίδευση των δικτύων καθώς και την ανάκληση της πληροφορίας που αυτή περιέχεται. Στόχος των αλγορίθμων αυτών είναι η προσομοίωση ευφυών διαδικασιών όπως είναι η αναγνώριση προτύπων, φυσικής γλώσσας αλλά και προσώπων, η πλοήγηση ρομπότ σε φυσικά εμπόδια, η ανάπτυξη βέλτιστων στρατηγικών για την επίλυση προβλημάτων και η εκτέλεση συλλογισμών για την εξαγωγή λογικών συμπερασμάτων. Αξίζει να σημειωθεί ότι στοχεύουν επιπλέον, στην αυτοπροσαρμογή σε νέες καταστάσεις και γνωστά περιβάλλοντα, καθώς και στην μάθηση από την εμπειρία τους. Προκειμένου να επιτευχθούν τα 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, την κατηγορία που χρησιμοποιείται στα περισσότερα δίκτυα. 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 παραπάνω, σε αυτή την κατηγορία μάθησης υπάρχει αντιστοίχιση ανάμεσα στην είσοδο και στην έξοδο. Έτσι, το δίκτυο συνέχεια επεξεργάζεται τις εισόδους με τα αντίστοιχα συναπτικά βάρη και συγκρίνει το αποτέλεσμα αυτό, 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, καθώς τα βάρη βελτιώνονται συνεχώς. Το σύστημα θεωρούμε πως έχει εκπαιδευτεί με επιτυχία όταν θα έχουμε μία ολόκληρη εποχή χωρίς να χρειάζεται αναπροσαρμογή των βαρών.

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. Αξίζει να τονίσουμε σε αυτό το σημείο πως 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. 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.[[9]](#footnote-8) Τέλος, σύμφωνα με τον Κίτρου, η μάθηση με επίβλεψη χρησιμοποιείται και σε προβλήματα πρόγνωσης αλλά και διερμηνείας.

## 1.6 Εφαρμογές νευρωνικών δικτύων

Τα ΤΝΔ είναι ιδιαίτερα δημοφιλή σε προβλήματα που περιέχουν μη-προβλέψιμες λειτουργίες. Μία από τις κατηγορίες προβλημάτων στις οποίες μπορούν να ενταχθούν οι εφαρμογές των ΤΝΔ είναι τα προβλήματα κατηγοριοποίησης (classification). Στον ιατρικό τομέα, αξιοποιούνται για τη διάγνωση παθήσεων και την ανάπτυξη νέων φαρμάκων για ασθένειες, επιταχύνοντας σημαντικά τη χρονοβόρα διαδικασία ελέγχου των δεδομένων ιατρικών εξετάσεων. Ωστόσο, η αναγκαιότητα των ιατρών παραμένει αναλλοίωτη. Στον τομέα της άμυνας, χρησιμοποιούνται για την κατηγοριοποίηση εικόνων προερχόμενων από συσκευές radar ή sonar. Στον επιχειρηματικό τομέα, εφαρμόζονται για την κατηγοριοποίηση των πελατών βάσει των αγοραστικών τους συνηθειών.

Τα προβλήματα αναγνώρισης (recognition/identification) είναι μία ακόμη κατηγορία προβλημάτων στις οποίες μπορούν να ενταχθούν οι εφαρμογές των ΤΝΔ. Στον τραπεζικό τομέα, ενσωματώνονται για τον έλεγχο της γνησιότητας των υπογραφών αλλά και των τραπεζογραμματίων. Στον τομέα της πληροφορικής, υιοθετούνται για την αναγνώριση ήχου -ακόμη και φωνής, εικόνας καθώς και γραπτού κειμένου είτε αυτό είναι χειρόγραφο είτε τυπωμένο.

Τα προβλήματα αποτίμησης (assessment) είναι μία ακόμη κατηγορία προβλημάτων στις οποίες μπορούν να ενταχθούν οι εφαρμογές των ΤΝΔ. Στον τομέα της άμυνας υπάρχει η παρακολούθηση στόχων. Στον τομέα της ασφάλειας εφαρμόζονται για τον εντοπισμό κίνησης, την ταύτιση δακτυλικών αποτυπωμάτων καθώς και την ανάλυση εικόνας σε συστήματα επιτήρησης. Στην μηχανολογία χρησιμοποιούνται εφαρμογές παρακολούθησης, επιθεώρησης και ελέγχου των προϊόντων.

Τέλος, στον τομέα της πρόβλεψης, ενσωματώνονται εφαρμογές πρόβλεψης ισοτιμίας νομισμάτων και τιμών μετοχών, καθώς και πρόβλεψης πωλήσεων, και χρηματιστηριακές προβλέψεις. Στη γεωργία γίνεται πρόβλεψη παραγωγής κυρίως με χρήση δορυφορικών εικόνων, ενώ στη μετεωρολογία, γίνεται πρόβλεψη του καιρού. Επίσης μπορούμε να συμπεριλάβουμε την αυτόματη πλοήγηση στο χώρο καθώς απαιτεί πρόβλεψη της κίνησης στο χώρο και δυναμικών εμποδίων. Μπορούμε να αναφέρουμε επίσης, τη λήψη αποφάσεων διότι περιλαμβάνει την πρόβλεψη των συνεπειών διάφορων επιλογών. Τέλος, η πρόβλεψη οικονομικών μεγεθών (πτώχευση, πωλήσεις) καθώς και η δυναμική μοντελοποίηση συστημάτων.

Άλλες κατηγορίες προβλημάτων στα οποία έχουν εφαρμογή τα ΤΝΔ είναι η προσέγγιση συνάρτησης, οι χρονοσειρές, η εξόρυξη δεδομένων και η όραση. Τα τελευταία χρόνια τα ΤΝΔ χρησιμοποιούνται σε συστήματα ελέγχου που βασίζονται στην ασαφή λογική (neuofuzzy systems) με κύριο ρόλο τον υπολογισμό της συνάρτησης συγγένειας.

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 υπάρχει κάποιο ΤΝΔ με σκοπό την αντιμετώπιση ετερογενών προβλημάτων. Υπάρχουν όμως προγραμματιστικά περιβάλλοντα τα οποία επιτρέπουν τη δόμηση ΤΝΔ για διάφορους σκοπούς. Αυτό είναι πολύ σημαντικό καθώς στο ερώτημα “Ποιο είδος ΤΝΔ είναι κατάλληλο για δεδομένο πρόβλημα” η απάντηση δεν είναι ούτε εύκολη, ούτε μοναδική. Κάθε πρόβλημα είναι συνήθως ειδική περίπτωση και η αντιμετώπιση του με ΤΝΔ απαιτεί μελέτη και πειραματισμό.[[10]](#footnote-9)

# Δομή και Λειτουργία των Τεχνητών Νευρωνικών Δικτύων

## Τα δίκτυα Perceptron

Τα δίκτυα Perceptron καταλαμβάνουν μία ειδική θέση στην ιστορική εξέλιξη των ΤΝΔ καθώς ήταν η πρώτη μορφή δικτύου που μπορούσε να περιγραφεί αλγοριθμικά. Χρησιμοποιούνται για την ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, αρκεί αυτά να ανήκουν σε δύο μόνο κατηγορίες. Η σπουδαιότητα τους δεν είναι μόνο ιστορική καθώς συνεχίζουν να έχουν πρακτική χρησιμότητα. (Haykin, 2010)

Τα δίκτυα αυτά έχουν αναπτυχθεί βάσει τριών (3) ερωτημάτων. Το πρώτο από αυτά είναι το πως γίνεται αισθητή η ανίχνευση πληροφοριών για τον φυσικό κόσμο, από το βιολογικό σύστημα. Το δεύτερο ερώτημα, αφορά την μορφή με την οποία αποθηκεύονται ή απομνημονεύονται οι πληροφορίες και το τρίτο και τελευταίο ερώτημα, αφορά τον τρόπο με τον οποίο οι πληροφορίες που περιέχονται στην αποθήκευση ή στη μνήμη επηρεάζουν την αναγνώριση και τη συμπεριφορά. Στόχος των δικτύων αυτών, λοιπόν είναι η απεικόνιση μερικών θεμελιωδών ιδιοτήτων των ευφυών συστημάτων χωρίς όμως να είναι απαραίτητο να εμβαθύνουν σε ειδικές συνθήκες που μπορεί να αφορούν συγκεκριμένους βιολογικούς οργανισμούς, οι οποίες συχνά μπορεί να είναι και άγνωστες. (Rosenblatt, 1958)

Το στοιχειώδες Perceptron (elementary Perceptron) αποτελεί την πιο απλή μορφή των δικτύων αυτών και περιλαμβάνει έναν και μοναδικό τεχνητό νευρώνα. Έχει αποδειχθεί πως το αποτέλεσμα του Perceptron είναι το ίδιο, ανεξάρτητα από την συνάρτηση ενεργοποίησης που θα χρησιμοποιηθεί, αν και η πιο σύνηθες συνάρτηση είναι η βηματική.[[11]](#footnote-10) Συνήθως τα βάρη αρχικοποιούνται τυχαία, στο διάστημα 0 έως 1.

Τα δίκτυα Perceptron αποτελούν την πιο απλή μορφή δικτύου πρόσθιας τροφοδότησης χρησιμοποιώντας μάθηση με επίβλεψη. Ζητούμενο αυτού του δικτύου είναι η αυτόματη εκμάθηση των παραμέτρων του συστήματος - δηλαδή των βαρών - με στόχο την επίτευξη του επιθυμητού στόχου. Ο κλασικός κανόνας εκπαίδευσης Perceptron είναι γνωστός και ως κανόνας σταθερής αύξησης (fixed increment rule) και είναι επαναληπτικός. Πρακτικά, τα πρότυπα παρουσιάζονται με κυκλική σειρά στο δίκτυο επαναληπτικά. Σε περίπτωση που υπάρχει σφάλμα ταξινόμησης - δηλαδή η πραγματική έξοδος είναι διαφορετική από την επιθυμητή - μόνο τότε γίνεται η εκπαίδευση των βαρών. Με το που γίνει η εκπαίδευση, έχει αποδειχθεί πως το πρότυπο αυτό, την επόμενη φορά που θα χρησιμοποιηθεί είτε θα ταξινομηθεί σωστά είτε θα πλησιάζει ακόμη περισσότερο στην σωστή ταξινόμηση του. (Διαμαντάρας, 2007)

## 2.2 Τα δίκτυα Adaline

Ο όρος Adaline είναι ένα ακρώνυμο των λέξεων ADaptive LINear Element και μεταφράζεται ως Αυτοπροσαρμοζόμενο Γραμμικό Στοιχείο. Ο όρος αυτός εισήχθη από τον Windrow. Η διαφορά σε σχέση με τον κλασικό μη γραμμικό νευρώνα των McCulloch-Pitts είναι ότι η έξοδος αυτού του δικτύου αλλά και οι στόχοι του προαιρετικά, μπορούν να πάρουν συνεχείς αντί για διακριτές τιμές. Ένα πρόβλημα που προκύπτει όμως, είναι πως επειδή οι τιμές των στόχων δεν προέρχονται από ένα διακριτό σύνολο, δεν γίνεται να είναι προφανή. 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.[[12]](#footnote-11)

## 2.3 Τα δίκτυα MLP

Όπως αναφέρθηκε σε προηγούμενο υποκεφάλαιο, το δίκτυο Perceptron μπορεί να χρησιμοποιηθεί για τον διαχωρισμό προτύπων μόνο δύο κατηγοριών εφόσον αυτά είναι γραμμικά διαχωρίσιμα. Μια εξέλιξη του Perceptron είναι τα δίκτυα MLP, τα οποία χρησιμοποιούν περισσότερους από έναν νευρώνα, αίροντας τους παραπάνω περιορισμούς. Με τη χρήση της βηματικής συνάρτησης είτε με την απλή είτε με τη διπολική της μορφή, τα MLP δίκτυα έχουν τη δυνατότητα να υλοποιήσουν συναρτήσεις που δεν έχει τη δυνατότητα να υλοποιήσει ένα δίκτυο Perceptron. Ωστόσο, προτιμάται η σιγμοειδής συνάρτηση έναντι της βηματικής καθώς αποδεικνύεται πως με την χρήση της, τα δίκτυα μπορούν να προσεγγίσουν οποιαδήποτε ομαλή συνάρτηση, όσο κοντά επιθυμούμε. Αυτός είναι και ο λόγος που αυτά τα δίκτυα καλούνται και Universal Approximators (Καθολικοί Προσεγγιστές). Αξιοσημείωτο είναι το γεγονός πως αρκεί να χρησιμοποιηθούν μόνο δύο στρώματα νευρώνων πέρα από το στρώμα εισόδου. Ένα από τα πιο βασικά προβλήματα που δεν μπορεί να λύσει το δίκτυο Perceptron αλλά μπορεί το δίκτυο MLP είναι η υλοποίηση της πύλης XOR (eXclusive OR).[[13]](#footnote-12)

Μπορούμε να πούμε πως η θεωρητική ανάλυση αυτών των δικτύων είναι αρκετά δύσκολη υπόθεση. Αυτό οφείλεται λόγω των κρυφών επιπέδων που υπάρχουν σε αυτά καθώς και της μεγάλης διασυνδεσιμότητας τους. Στη γενική του μορφή το δίκτυο είναι πλήρως συνδεδεμένο (fully connected). Στα δίκτυα MLP μπορούμε να αναγνωρίσουμε δύο είδη σημάτων. Ένα από αυτά είναι τα λειτουργικά σήματα ή αλλιώς σήματα εισόδου. Όπως λέει και το όνομα τους, είναι σήματα που φτάνουν στην είσοδο του δικτύου, διαδίδονται σε όλο το δίκτυο και εν τέλει φτάνουν στην έξοδο του δικτύου σαν σήμα εξόδου. Το δεύτερο είδος σημάτων είναι τα σήματα σφάλματος (error signal) τα οποία προέρχονται από έναν νευρώνα εξόδου και διαδίδεται προς τα πίσω διαμέσου του δικτύου. Έτσι, κάθε νευρώνας εξόδου σχεδιάζεται με τέτοιο τρόπο έτσι ώστε να εκτελεί δύο υπολογισμούς. Αρχικά, τον υπολογισμό του λειτουργικού σήματος που εμφανίζεται στην έξοδο κάθε νευρώνα, ο οποίος εκφράζεται ως μία συνεχής μη γραμμική συνάρτηση του σήματος εισόδου και των συναπτικών βαρών που σχετίζονται με αυτόν τον νευρώνα. Ο δεύτερος υπολογισμός που γίνεται μέσω του νευρώνα εξόδου, είναι αυτός της εκτίμησης του διανύσματος κλίσης ο οποίος χρειάζεται για την κατάσταση κατά την οποία το δίκτυο εξελίσσεται προς τα πίσω. Οι κρυφοί νευρώνες, δρουν ως ανιχνευτές χαρακτηριστικών (feature detectors) παίζοντας κρίσιμο ρόλο στη λειτουργία των MLP δικτύων. Καθώς προχωράει η διαδικασία μάθησης, οι κρυφοί νευρώνες είναι αυτοί που σταδιακά “ανακαλύπτουν” τα ξεχωριστά χαρακτηριστικά των προτύπων, κάτι το οποίο διαφορεοποιεί τα MLP δίκτυα με τα δίκτυα Perceptron του Rosenblatt.[[14]](#footnote-13)

## 2.4 Η μέθοδος Back Propagation

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, σύμφωνα με την οποία - όπως αναφέρθηκε σε προηγούμενο υποκεφάλαιο, αν έχουμε το κατάλληλο σε μέγεθος δίκτυο μπορούμε να το εκπαιδεύσουμε να μάθει οποιαδήποτε συνάρτηση εμείς επιθυμούμε σε αντίθεση με τα δίκτυα Perceptron που μπορούν να κατηγοριοποίησουν μόνο γραμμικά διαχωρίσιμα πρότυπα. Κυριότερος εκπρόσωπος των αλγορίθμων εκπαίδευσης MLP είναι ο Back Propagation.

O αλγόριθμος Back-Propagation προτάθηκε τη 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. 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. (Διαμαντάρας)

Η κεντρική ιδέα λοιπόν της μεθόδου Back Propagation, είναι ο καθορισμός και η ελαχιστοποίηση του σφάλματος σε κάθε νευρώνα για κάθε στρώμα του ΤΝΔ. Η εκπαίδευση των βαρών γίνεται από το στρώμα εξόδου προς το στρώμα εισόδου.

Αξίζει να αναφερθεί βέβαια πως ο αλγόριθμος αυτός έχει αρκετά προβλήματα. Το βασικότερο αυτών είναι η αργή σύγκλιση Βέβαια, αναφέρεται πως δεν υπάρχει και εγγύηση γενίκευσης. (Δούνιας & Καραμπότσης, χ.χ.)

## 2.5 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).

Μπορεί να αποδοθεί ως επικλινής κάθοδος ή επικλινής κατάδυση. (Κουτρούμπας, θεωδωριδης)

# Γενετικοί αλγόριθμοι

## Ιστορική Αναδρομή

Οι γενετικοί αλγόριθμοι αναπτύχθηκαν από τον John Holland, τους συναδέλφους του αλλά και τους μαθητές του στο πανεπιστήμιο του Michigan. Προτάθηκαν τη δεκαετία του 1960 και πιο συγκεκριμένα, το 1962. Ο πρώτος στόχος της έρευνας του ήταν η αφαίρεση και η αυστηρή εξήγηση των προσαρμοστικών διαδικασιών των φυσικών συστημάτων και ο δεύτερος ήταν η σχεδίαση λογισμικού τεχνητών συστημάτων που διατηρεί τους σημαντικούς μηχανισμούς των φυσικών συστημάτων.

Το κεντρικό θέμα της έρευνας των γενετικών αλγορίθμων ήταν η ευρωστία. Βρίσκουν εφαρμογή σε επιχειρηματικούς, επιστημονικούς αλλά και μηχανικούς κύκλους.

Αποτελούν τους δημοφιλέστερους από τους εξελικτικούς αλγόριθμους και βασίζονται στις αρχές της εξέλιξης των ειδών που εισήγαγε ο Δαρβίνος με το βιβλίο “Η καταγωγή των ειδών”. Η αρχική ορολογία που εισήγαγε ο Holland ήταν γενετικό σχέδιο, αλλά ο όρος γενετικός αλγόριθμος υιοθετήθηκε αργότερα για να υπογραμμίσει τον κεντρικό ρόλο του υπολογισμού στον καθορισμό και την υλοποίηση των σχεδίων.

Οι γενετικοί αλγόριθμοι είναι αλγόριθμοι αναζήτησης που βασίζονται στη μηχανική της φυσικής επιλογής και γενετικής. Συνδυάζουν την επιβίωση του καταλληλότερου με μία δομημένη αλλά τυχαιοποιημένη ανταλλαγή πληροφοριών για να σχηματίσουν έναν αλγόριθμο αναζήτησης με κάποιο από τα καινοτόμα ίχνη της ανθρώπινης αναζήτησης. (Καμπουρλάζος & Παπακώστας, 2015)

## 3.2 Μέθοδοι κωδικοποίησης

Κυρίως κείμενο της εργασίας.

## 3.3 Γενετικοί τελεστές

Κυρίως κείμενο της εργασίας.

## 3.4 Παράλληλοι γενετικοί αλγόριθμοι

Κυρίως κείμενο της εργασίας.

## 3.5 Πλεονεκτήματα γενετικών αλγορίθμων

Η χρήση των γενετικών αλγορίθμων είναι ελκυστική για αρκετούς λόγους. Αρχκά, μπορούν να λύσουν δύσκολα προβλήματα γρήγορα αλλά και αξιόπιστα.Η αποδοτικότητα είναι ένας από τους σημαντικούς λόγους χρήσης τους. Τόσο η θεωρία όσο και η πράξη έχουν δείξει πως προβλήματα που έχουν πολλές, δύσκολα προσδιορισμένες λύσεις μπορούν να αντιμετωπιστούν καλύτερα από τους γενετικούς αλγόριθμους. Αξιοσημείωτο είναι πως για τους γενετικούς αλγορίθμους δεν αποτελούν σημεία δυσχέρειας, συναρτήσεις που παρουσιάζουν μεγάλες διακυμάνσεις ενώ αυτές καθιστούν ανεπαρκείς άλλες μεθόδους στην εύρεση των ακρότατων τους.

Ένα ακόμη πλεονέκτημα των γενετικών αλγορίθμων είναι πως μπορούν εύκολα να συνεργαστούν με τα υπάρχοντα μοντέλα και συστήματα χωρίς να χρειάζεται η επανασχεδίαση τους. Αυτό οφείλεται στο ότι χρησιμοποιούν μόνο πληροφορίες της διαδικασίας ή συνάρτησης που πρόκειται να βελτιστοποιήσουν. Επιπλέον, είναι εύκολα επεκτάσιμοι και εξελίξιμοι χωρίς να είναι απαραίτητο πως οι λειτουργίες τους πρέπει να είναι δανεισμένες από τη φύση.

Οι γενετικοί τελεστές έχουν τη δυνατότητα να συμμετέχουν σε υβριδικές μορφές με άλλες μεθόδους. Παρόλο που η ισχύς τους είναι μεγάλη, σε μερικές ειδικές περιπτώσεις προβλημάτων, όπου άλλες μέθοδοι συμβαίνει να έχουν πολύ υψηλή αποδοτικότητα λόγω εξειδίκευσης, υπάρχει η δυνατότητα χρησιμοποίησης ενός υβριδικού σχήματος γενετικών αλγορίθμων με άλλη μέθοδο και αυτό δίνει μεγάλη ευελιξία σε αυτούς τους αλγορίθμους.

Παρόλο που η εξερεύνηση του χώρου και η εκμετάλλευση της επεξεργασμένης πληροφορίας είναι ανταγωνιστικά χαρακτηριστικά, οι γενετικοί αλγόριθμοι τα συνδυάζουν, κάτι που τους κάνει αρκετά αποδοτικούς αλλά και ελκυστικούς. Ένα ακόμη στοιχείο που εκμεταλλεύονται οι γενετικοί αλγόριθμοι είναι ότι επιδέχονται παράλληλη υλοποίηση, κάτι που αυξάνει ακόμη περισσότερο την απόδοση τους.

Ένας από τους κύριους λόγους που καθιστά τις παραδοσιακές μεθόδους δύσκαμπτες και ακατάλληλες για πολλά προβλήματα είναι η απαίτηση τους για την ύπαρξη περιορισμών, κάτι το οποίο είναι αδιάφορο για τους γενετικούς αλγορίθμους.Έτσι, μπορούν να χρησιμοποιούν σε πάρα πολλούς τομείς όπως στην οικονομία, στον σχεδιασμό μηχανών, στην επίλυση μαθηματικών εξισώσεων, στην εκπαίδευση ΤΝΔ και σε πολλούς άλλους τομείς.(Λυκοθανάσης & Κουτσομητρόπουλος, 2021)

## 3.6 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)

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.

# 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

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

Γεωργούλη, Κ. (2015) *Τεχνητή Νοημοσύνη - Μία εισαγωγική προσέγγιση*. Αθήνα: Κάλλιπος

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

Δούνιας, Γ. και Καραμπότσης, Ε. (χ.χ.) *Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα.* Διαθέσιμο σε: [URL](https://mde-lab.aegean.gr/images/stories/docs/ANNs-intro.pdf) (Ημερομηνία πρόσβασης: 15 Ιουνίου 2024).

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

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

Λύκας, Α. (χ.χ.) *Βασικές αρχές εκπαίδευσης ΤΝΔ: Το perceptron*. Διαθέσιμο σε: [URL](https://www.cs.uoi.gr/~arly/courses/nn/slides/K2.pdf) (Ημερομηνία πρόσβασης: 10 Ιουνίου 2024).

Λυκοθανάσης, Σ., Κουτσομητρόπουλος, Δ. (2021) *Υπολογιστική νοημοσύνη και βαθιά μάθηση*. Εκδόσεις Κάλλιπος

Ντούνης, Α. (2024) *Βασικές Αρχές Υπολογιστικής Νοημοσύνης*. Εκδόσεις Κάλλιπος

Πούλος, Μ. (2015) *Σημασιολογική Επεξεργασία της Πληροφορίας*. Εκδόσεις Κάλλιπος

Τσίπουρας, Μ. (2015-2016) *Τεχνητή Νοημοσύνη.*Διαθέσιμο σε: [URL](https://eclass.teiwm.gr/modules/document/file.php/DNG170/%CE%94%CE%B9%CE%B1%CE%BB%CE%AD%CE%BE%CE%B5%CE%B9%CF%82/%CE%9C%CE%AC%CE%B8%CE%B7%CE%BC%CE%B1%2010%20-%20%CE%9D%CE%B5%CF%85%CF%81%CF%8E%CE%BD%CE%B1%CF%82%20Perceptron.pdf) (Ημερομηνία πρόσβασης: 15 Ιουνίου 2024).

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).

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.

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

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) & (Λυκοθανάσης & Κουτσημητρόπουλος, 2021) & (Γεωργούλη, 2015) & (Ντούνης, 2024) & (Haykin, 2010) & (Τσίπουρας, 2015) [↑](#footnote-ref-0)
2. (Δούνιας & Καραμπότσης, χ.χ.) [↑](#footnote-ref-1)
3. (Διαμαντάρας, 2007) & (Βλαχάβας & Κεφάλας & Βασιλειάδης & Κόκκορας & Σακελλαρίου, 2006) & (Haykin, 2010) [↑](#footnote-ref-2)
4. (Βλαχάβας et al., 2006) & (Haykin, 2010) & (Πούλος, 2015) [↑](#footnote-ref-3)
5. Χρησιμοποιείται στο δίκτυο Adaline (Δούνιας & Καραμπότσης, χ.χ.) [↑](#footnote-ref-4)
6. Αλλιώς αναφέρεται και ως συνάρτηση κατωφλίου (Λύκας, χ.χ.) Χρησιμοποιείται στον νευρώνα McCullich & Pitts, αλλά και στο δίκτυο Perceptron (Δούνιας & Καραμπότσης, χ.χ.) Δεν προτιμάται σε δίκτυο MLP. Είναι η πιο απλή συνάρτηση ενεργοποίησης και για αυτόν τον λόγο, χρησιμοποιείται περισσότερο σε θεωρητικές μελέτες. (Chiarandini, χ.χ.) & (Τσούλος, χ.χ.) Η έξοδος είναι ένας δυαδικός αριθμός είτε με την κλασική μορφή (0/1) είτε με τη λεγόμενη διπολική μορφή (-1/1). Η επιλογή ωστόσο μεταξύ αυτών των μορφών είναι ήσσονος σημασίας (Διαμαντάρας, 2007) [↑](#footnote-ref-5)
7. Χρησιμοποιείται περισσότερο σε πρακτικές εφαρμογές. (Chiarandini, χ.χ.) Η γραφική της παράσταση έχει το σχήμα “S” και είναι η πιο κοινή μορφή συνάρτησης ενεργοποίησης που χρησιμοποιείται για την κατασκευή ΤΝΔ. (Haykin, 2010) & (Διαμαντάρας, 2007) [↑](#footnote-ref-6)
8. (Βλαχάβας et al., 2006) & (Λύκας, 2008) [↑](#footnote-ref-7)
9. (Haykin, 2010) & (Θεοδωρίδης & Κουτρούμπας, 2012) & (Κίτρου, χ.χ.) & (Διαμαντάρας, 2007) & (Λυκοθανάσης, 2021) [↑](#footnote-ref-8)
10. (Τσούλος, χ.χ.) & (Βλαχάβας et al., 2006) & (Ντούνης, 2024) [↑](#footnote-ref-9)
11. (Βλαχάβας et al., 2006) & (Haykin, 2010) [↑](#footnote-ref-10)
12. (Διαμαντάρας, 2007) [↑](#footnote-ref-11)
13. (Διαμαντάρας, 2007) [↑](#footnote-ref-12)
14. (Haykin, 2010) [↑](#footnote-ref-13)