

**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.

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](#_Toc7972)

[ΠΕΡΙΛΗΨΗ 10](#_Toc9139)

[ABSTRACT 11](#_Toc16994)

[ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ 12](#_Toc17457)

[ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ 14](#_Toc15454)

[ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ 15](#_Toc21876)

[ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ 16](#_Toc22030)

[ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ 17](#_Toc16030)

[1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα 18](#_Toc7167)

[1.1 Ιστορική Αναδρομή 18](#_Toc31771)

[1.2 Μετάβαση από τους βιολογικούς στους τεχνητούς νευρώνες 20](#_Toc31421)

[1.3 Συναρτήσεις ενεργοποίησης 23](#_Toc19175)

[1.4 Τεχνητά Νευρωνικά Δίκτυα: Μία σφαιρική εικόνα 26](#_Toc15210)

[1.5 Μάθηση με επίβλεψη 28](#_Toc14158)

[1.6 Εφαρμογές νευρωνικών δικτύων 30](#_Toc18958)

[2. Δομή και Λειτουργία των Τεχνητών Νευρωνικών Δικτύων 32](#_Toc29303)

[2.1 Τα δίκτυα Perceptron 32](#_Toc11640)

[2.2 Τα δίκτυα Adaline 34](#_Toc30286)

[2.3 Τα δίκτυα MLP 35](#_Toc28093)

[2.4 Η μέθοδος Back Propagation 38](#_Toc17609)

[2.5 Η μέθοδος Gradient Descent 41](#_Toc11792)

[3. Γενετικοί αλγόριθμοι 42](#_Toc29607)

[3.1. Ιστορική Αναδρομή 51](#_Toc30858)

[3.2 Μέθοδοι κωδικοποίησης 52](#_Toc6942)

[3.3 Γενετικοί τελεστές 53](#_Toc18719)

[3.4 Παράλληλοι γενετικοί αλγόριθμοι 54](#_Toc16788)

[3.5 Πλεονεκτήματα γενετικών αλγορίθμων 55](#_Toc17850)

[3.6 Εφαρμογές γενετικών αλγορίθμων 60](#_Toc20966)

[4. Μέθοδος - αποτελέσματα 61](#_Toc32764)

[4.1 Τα dataset που χρησιμοποιήθηκαν 61](#_Toc26776)

[4.2 Η προτεινόμενη μέθοδος 62](#_Toc22038)

[4.3 Πειραματικά αποτελέσματα 63](#_Toc22486)

[5. Συμπεράσματα 64](#_Toc29530)

[ΠΑΡΑΡΤΗΜΑ 65](#_Toc48)

[ΒΙΒΛΙΟΓΡΑΦΙΑ 70](#_Toc13945)

# 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

OCR………………......................................................….Οπτική Αναγνώριση Χαρακτήρων

# ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ

**ΠΡΟΤΥΠΑ:**

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

# 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 πρώτη δημοσίευση από τους Αμερικανούς επιστήμονες McCulloch και Pitts. Η δημοσίευση αυτή αφορούσε την ανάπτυξη ενός μοντέλου νευρωνικού δικτύου, το οποίο έκτοτε 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.

Δεκαπέντε χρόνια αργότερα, το 1958, ο ψυχολόγος Rosenblatt δημοσίευσε το δίκτυο Perceptron, μία καινοτόμα μέθοδο μάθησης με επίβλεψη. Προτάθηκε ως ένας μηχανισμός που μπορεί να εκπαιδευτεί στην κατηγοριοποίηση προτύπων και σε διάφορες παραλλαγές, πιο απλές ή ακόμη και πιο σύνθετες, εξακολουθεί να υφίσταται μέχρι και σήμερα. Το δίκτυο Perceptron, αποτελεί ένα απλό μοντέλο δύο επιπέδων και είναι μία ελαφρώς τροποποιημένη έκδοση του νευρώνα των McCulloch και Pitts. 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 ακόμη και ως προς την έννοια της δυαδικής ισοτιμίας. 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 και αποτελεί τροποποίηση του δικτύου Perceptron. Αποτελεί 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. Οι συνάψεις 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 σύνδεση 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 παράγει πολύ αραιά παλμούς σε τυχαίες στιγμές 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 στις περισσότερες περιπτώσεις, 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

Τι είναι τελικά τα ΤΝΔ; Είναι συστήματα επεξεργασίας δεδομένων που αποτελούνται από ένα πλήθος τεχνητών νευρώνων οργανωμένων σε δομές παρόμοιες με αυτές του ανθρώπινου εγκεφάλου. Συνήθως αυτοί οι νευρώνες 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. Ουσιαστικά σε αυτό το επίπεδο δεν γίνεται κάποια επεξεργασία και για αυτόν τον λόγο δεν είναι ουσιαστικά νευρώνες. Στη συνέχεια, υπάρχουν προαιρετικά ένα ή περισσότερα κρυφά επίπεδα (hidden layer) ενώ στο τέλος υπάρχει το επίπεδο εξόδου (output layer).

Οι νευρώνες των ΤΝΔ 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 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. Η ανάκληση της πληροφορίας γίνεται βάσει του περιεχομένου και όχι της διεύθυνσης, όπως συμβαίνει στον ανθρώπινο εγκέφαλο. H τρίτη 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 είτε λόγω θορύβου είτε επειδή 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) καθώς δεν επηρεάζονται από ελλιπή ή/και δεδομένα με θόρυβο. 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 διαστημικές αποστολές, σε χώρους με ραδιενέργεια, ακόμη και σε πεδία μάχης. (Βλαχάβας et al., 2006)

## 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, 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. 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 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.[[8]](#footnote-7) 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. Υπάρχουν διαφορετικές κατηγορίες προβλημάτων στις οποίες μπορούν να ενταχθούν οι εφαρμογές των ΤΝΔ.

Αρχικά θα αναφερθούμε στα προβλήματα κατηγοριοποίησης (classification), αναφέροντας πρώτο τον ιατρικό τομέα με εφαρμογές όπως είναι η διάγνωση παθήσεων και η ανάπτυξη νέων φαρμάκων για ασθένειες. Μέσα από αυτές τις εφαρμογές επιταχύνεται σημαντικά η χρονοβόρα διαδικασία ελέγχου των δεδομένων ιατρικών εξετάσεων από ιατρούς. Ωστόσο αυτό δεν συνεπάγεται τη μη χρησιμότητα των ιατρών. Στον τομέα της άμυνας, γίνεται η κατηγοριοποίηση εικόνων προερχόμενων από συσκευές radar ή sonar. Στις επιχειρήσεις, γίνεται κατηγοριοποίηση των πελατών βάση των αγοραστικών τους συνηθειών. Τέλος, γίνεται και η ανίχνευση βλαβών σε μηχανήματα.

Επίσης, υπάρχουν τα προβλήματα αναγνώρισης (recognition/identification), στα οποία 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 αλλά και φωνής. Εδώ εντάσσονται και οι πιο διάσημες εφαρμογές των ΤΝΔ, οι εφαρμογές OCR.

Σε προβλήματα αποτίμησης (assessment), μπορούμε να αναφέρουμε την παρακολούθηση στόχων στον τομέα της άμυνας, τον εντοπισμό κίνησης, την ταύτιση δακτυλικών αποτυπωμάτων καθώς και την ανάλυση εικόνας σε συστήματα επιτήρησης στον τομέα της ασφάλειας. Επιπλέον, στη μηχανολογία έχουμε εφαρμογές παρακολούθησης, επιθεώρησης και ελέγχου των προϊόντων. Τέλος μπορούμε να συμπεριλάβουμε και την ανάπτυξη επιχειρημάτων καθώς χρειάζεται αποτίμηση των δεδομένων για την υποστήριξη ενός επιχειρήματος.

Τέλος, στον τομέα της πρόβλεψης, έχουμε 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 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 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.[[9]](#footnote-8)

# Δομή και Λειτουργία των Τεχνητών Νευρωνικών Δικτύων

## Τα δίκτυα Perceptron

Το δίκτυο Perceptron καταλαμβάνει μία ειδική θέση στην ιστορική εξέλιξη των ΤΝΔ καθώς ήταν το πρώτο δίκτυο που μπορούσε να περιγραφεί αλγοριθμικά. Αποτελεί την απλούστερη δυνατή μορφή ενός ΤΝΔ που χρησιμοποιείται για την ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, αρκεί αυτά να ανήκουν σε δύο μόνο κατηγορίες. Το Perceptron είναι ένα ΤΝΔ που αποτελείται από έναν μεμονωμένο νευρώνα και η λειτουργία του είναι βασισμένη στη μέθοδο μάθησης με συσχετισμό σφάλματος. Έχει αποδειχθεί πως το αποτέλεσμα του Perceptron είναι το ίδιο, ανεξάρτητα από την συνάρτηση ενεργοποίησης που θα χρησιμοποιηθεί. Η σπουδαιότητα του δεν είναι μόνο ιστορική καθώς συνεχίζει να έχει πρακτική χρησιμότητα. Η απόδειξη της σύγκλισης του αλγορίθμου είναι γνωστή ως θεώρημα σύγκλισης του Perceptron.[[10]](#footnote-9)

Ο Perceptron έχει αναπτυχθεί βασισμένος σε τρία (3) ερωτήματα. Το πρώτο από αυτό 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. 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 Perceptron είναι η απεικόνιση μερικών θεμελιωδών ιδιοτήτων των ευφυών συστημάτων, χωρίς όμως να εμπλέκεται πολύ βαθιά σε ειδικές συνθήκες που μπορεί να ισχύουν για συγκεκριμένους βιολογικούς οργανισμούς, οι οποίες συχνά μπορεί να είναι και άγνωστες.[[11]](#footnote-10)

Τα δίκτυα Perceptron είναι η πιο απλή μορφή δικτύου πρόσθιας τροφοδότησης, χωρίς κρυφά επίπεδα, χρησιμοποιώντας μάθηση με επίβλεψη. Ζητούμενο αυτού του δικτύου είναι η αυτόματη εκμάθηση των παραμέτρων του συστήματος, 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. Στην πράξη, τα πρότυπα παρουσιάζονται με κυκλική σειρά στο δίκτυο και όταν τελειώσουν επαναλαμβάνονται από την αρχή. Ένας πλήρης κύκλος χρήσης όλων των προτύπων καλείται εποχή (epoch). Ο κανόνας τροποποιεί το διάνυσμα των συναπτικών βαρών μόνο όταν υπάρχει σφάλμα ταξινόμησης, δηλαδή όταν υπάρχει διαφορά ανάμεσα στην πραγματική και την επιθυμητή έξοδο. Έχει αποδειχτεί πως κάθε φορά που γίνεται η εκπαίδευση, γίνεται με τέτοιο τρόπο έτσι ώστε το πρότυπο που ταξινομήθηκε εσφαλμένα, στην επόμενη επανάληψη είτε θα ταξινομηθεί σωστά είτε θα πλησιάζει περισσότερο στο να ταξινομηθεί σωστά. Αποδεικνύεται πως το μοντέλο Perceptron συγκλίνει σε μία λύση, σε πεπερασμένο αριθμό επαναλήψεων, αρκεί το πρόβλημα να είναι γραμμικά διαχωρίσιμο. Αν όμως δεν είναι, τότε ο αλγόριθμος δεν συγκλίνει ποτέ κάτι το οποίο είναι σοβαρό μειονέκτημα και αποτέλεσε βασικό σημείο κριτικής εναντίον του Perceptron. Το κακό είναι πως τα περισσότερα προβλήματα στον πραγματικό κόσμο είναι μη γραμμικά διαχωρίσιμα.[[12]](#footnote-11)

Η πιο απλή μορφή του είναι το στοιχειώδες Perceptron (elementary Perceptron), το οποίο περιλαμβάνει έναν και μοναδικό τεχνητό νευρώνα χρησιμοποιώντας τη βηματική συνάρτηση ως συνάρτηση ενεργοποίησης. Η μάθηση είναι καθοδηγούμενη από το σφάλμα (error driven) και χρησιμοποιείται προκειμένου να εκπαιδευτούν κατάλληλα οι τιμές των βαρών έτσι ώστε να υπάρχει αντιστοιχία ανάμεσα στην πραγματική και επιθυμητή έξοδο. Συνήθως τα βάρη αρχικοποιούνται τυχαία, στο διάστημα 0 έως 1. [[13]](#footnote-12)

## 2.2 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 τελείως η μη γραμμική συνάρτηση. Η διαφορά σε σχέση με τον κλασικό μη γραμμικό νευρώνα των McCulloch-Pitts είναι ότι η έξοδος στο Adaline δίκτυο παίρνει συνεχείς τιμές και όχι διακριτές.Προεραιτικά, ακόμη και οι στόχοι μπορεί 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 διακριτό 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.[[14]](#footnote-13)

## 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. Υπάρχουν όμως δίκτυα που χρησιμοποιούν παραπάνω από έναν νευρώνα και αίρουν τους παραπάνω περιορισμούς. Αυτά καλούνται δίκτυα MLP. Σε περίπτωση που χρησιμοποιηθεί η βηματική συνάρτηση 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 για τα δίκτυα MLP και αποδεικνύεται πως με την χρήση αυτής της συνάρτησης, 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).[[15]](#footnote-14)

Οι ακόλουθες παρατηρήσεις υπογραμμίζουν τα βασικά χαρακτηριστικά των δικτύων MLP. Αρχικά, το μοντέλο κάθε νευρώνα στο δίκτυο περιλαμβάνει μια μη γραμμική συνάρτηση ενεργοποίησης. Δεύτερον, το δίκτυο περιέχει ένα ή περισσότερα επίπεδα τα οποία παραμένουν κρυφά (hidden) για τους κόμβους των επιπέδων εισόδου και εξόδου. Τέλος, το δίκτυο επιδεικνύει μεγάλη διασυνδεσιμότητα, ο βαθμός της οποίας καθορίζεται από τα συναπτικά βάρη του δικτύου. Αυτά τα χαρακτηριστικά, είναι επίσης υπεύθυνα για την ανεπάρκεια της γνώσης μας αναφορικά με τη συμπεριφορά του δικτύου. Πρώτον, η παρουσία μιας κατανεμημένης μορφής μη γραμμικότητας και η υψηλή διασυνδεσιμότητα του δικτύου καθιστούν τη θεωρητική ανάλυση ενός MLP δικτύου, δύσκολη υπόθεση. Δεύτερον, η χρήση κρυφών νευρώνων καθιστά δυσκολότερη την κατανόηση και σχηματοποίηση της διαδικασίας μάθησης. Εμμέσως, η διαδικασία μάθησης οφείλει να αποφασίσει ποια χαρακτηριστικά του προτύπου εισόδου θα πρέπει να αναπαρίσταται από τους κρυφούς νευρώνες. Έτσι, η διαδικασία μάθησης γίνεται πιο δύσκολη, επειδή η αναζήτηση πρέπει να διεξαχθεί σε έναν πολύ μεγαλύτερο χώρο πιθανών λειτουργιών και πρέπει να κάνει μια επιλογή μεταξύ των εναλλακτικών αναπαραστάσεων του προτύπου εισόδου. Στη γενική μορφή του το δίκτυο είναι πλήρως συνδεδεμένο (fully connected). Επομένως, ένας νευρώνας σε οποιοδήποτε επίπεδο του δικτύου συνδέεται με όλους τους νευρώνες του προηγούμενου επιπέδου. Η ροή σήματος διαμέσου του δικτύου προχωρά με κατεύθυνση προς τα εμπρός, από τα αριστερά προς τα δεξιά και επίπεδο προς επίπεδο. Στο δίκτυο MLP μπορούμε να αναγνωρίσουμε δύο είδη σημάτων, τα λειτουργικά σήματα και τα σήματα σφάλματος. Ένα λειτουργικό σήμα ή αλλιώς σήμα εισόδου είναι ένα σήμα το οποίο φτάνει στην είσοδο του δικτύου, διαδίδεται προς τα εμπρός (νευρώνα προς νευρώνα) σε όλο το δίκτυο και εμφανίζεται τελικά στην έξοδο του δικτύου, σαν σήμα εξόδου. Ένα σήμα σφάλματος (error signal) προέρχεται από ένα νευρώνα εξόδου του δικτύου και διαδίδεται προς τα πίσω (επίπεδο προς επίπεδο) διαμέσου του δικτύου. Οι νευρώνες εξόδου απαρτίζουν το επίπεδο εξόδου του δικτύου. Οι υπόλοιποι νευρώνες, απαρτίζουν τα κρυφά επίπεδα του δικτύου. Οι έξοδοι που παράγονται από το πρώτο κρυφό επίπεδο εφαρμόζονται στο επόμενο κρυφό επίπεδο, η διαδικασία συνεχίζεται κατά τον ίδιο τρόπο για όλο το υπόλοιπο του δικτύου. Κάθε κρυφός νευρώνας ή νευρώνας εξόδου ενός MLP σχεδιάζεται ώστε να εκτελεί δύο υπολογισμούς. Πρώτον, τον υπολογισμό του λειτουργικού σήματος που εμφανίζεται στην έξοδο κάθε νευρώνα, ο οποίος εκφράζεται ως μια συνεχής μη γραμμική συνάρτηση του σήματος εισόδου και των συναπτικών βαρών που σχετίζονται με αυτό τον νευρώνα. Δεύτερον, τον υπολογισμό μιας εκτίμησης του διανύσματος κλίσης (δηλαδή των κλίσεων της επιφάνειας σφάλματος σε σχέση με τα βάρη που είναι συνδεδεμένα στις εισόδους ενός νευρώνα), ο οποίος χρειάζεται για τη φάση που εξελίσσεται προς τα πίσω στο δίκτυο. Οι κρυφοί νευρώνες δρουν ως ανιχνευτές χαρακτηριστκών (feature detectors). Ως τέτοιοι, παίζουν κρίσιμο ρόλο στη λειτουργία ενός MLP. Καθώς προχωρά η διαδικασία μάθησης, οι κρυφοί νευρώνες αρχίζουν σταδιακά να “ανακαλύπτουν” τα εξέχοντα χαρακτηριστικά - αυτά που χαρακτηρίζουν τα δεδομένα εκπαίδευσης. Αυτό το κάνουν εκτελώντας ένα μη γραμμικό μετασχηματισμό στα δεδομένα εισόδου, σε ένα νέο χώρο που αποκαλείται χώρος χαρακτηριστικών. Σε αυτό τον νέο χώρο, οι κλάσεις χαρακτηριστικών που ενδιαφέρουν μια εργασία ταξινόμησης προτύπων, για παράδειγμα, μπορούν να ξεχωρίσουν από ό,τι άλλο θα μπορούσε να υπάρχει στον αρχικό χώρο δεδομένων εισόδου. Ο σχηματισμός αυτού του χώρου χαρακτηριστικών μέσω της επιβλεπόμενης μάθησης είναι που διαφοροποιεί το MLP από το Perceptron του Rosenblatt.[[16]](#footnote-15)

## 2.4 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.

(Δούνιας & Καραμπότσης, χ.χ.) - Αλγοριθμος Οπισθοδιάδοσης του  
Σφάλματος

* Βασίζεται στον κανόνα δ για συνεχείς συναρτήσεις ενεργοποίησης
* Δημοφιλής για τα πολυ-επίπεδα ΤΝΔ
* Κεντρική ιδέα: Καθορισμός και ελαχιστοποίηση του σφάλματος σε κάθε νευρώνα για κάθε στρώμα του ΤΝΔ.
* Η αναπροσαρμογή των βαρών γίνεται από το στρώμα εξόδου προς το στρώμα εισόδου.
* Ακολουθείται η διαδικασία βελτιστοποίησης επικλινούς καθόδου
* Η απόδοση του δικτύου πραγματοποιείται βάση προτύπων αξιολόγησης (Validation Set)

Προβλήματα:

* Παρουσιάζονται τοπικά ελάχιστα
* Ένα ή περισσότερα βάρη έχουν σταθερά υψηλές απόλυτες τιμές και δεν τροποποιούνται σημαντικά σε κάθε διόρθωση (Network paralysis)
* Δε μπορεί να υπάρξει εγγύηση γενίκευσης, ακόμη και με ελάχιστον σφάλμα δ

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.

## 2.5 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

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)

(Λυκοθανάσης)

Τα τελευταία τριάντα χρόνια έχει παρατηρηθεί ένα συνεχώς αυξανόμενο ενδιαφέρον για ανάπτυξη συστημάτων

επίλυσης προβλημάτων βασισμένων στις αρχές της *Γενετικής Εξέλιξης* και της *Κληρονομικότητας*. Τα

μειονεκτήματα των κλασικών μεθόδων αναζήτησης και βελτιστοποίησης, καθώς και η διαρκώς αυξανομένη

ανάγκη για παραγωγή λογισμικού που να μπορεί να εκμεταλλεύεται πιο αποδοτικά τις τεράστιες δυνατότητες

του υλικού, ήταν η βασική αιτία που ώθησε τους επιστήμονες σε αυτήν την αναζήτηση. Αυτού του είδους τα

συστήματα λειτουργούν διατηρώντας έναν πληθυσμό κωδικοποιημένων πιθανών λύσεων και εφαρμόζοντας

πάνω του διάφορες διαδικασίες επιλογής του καλύτερου, καθώς και διάφορους *γενετικούς τελεστές*. Οι

συγκεκριμένοι τελεστές αντιγράφουν τον τρόπο με τον οποίο αναπαράγονται και μεταλλάσσονται τα

χρωμοσώματα των κυττάρων των ζωντανών οργανισμών. Έτσι, περνώντας από γενιά σε γενιά, τα συστήματα

αυτά δημιουργούν συνεχώς νέους πληθυσμούς πιθανών λύσεων χρησιμοποιώντας τόσο κομμάτια και στοιχεία 18

Σπυρίδων Λυκοθανάσης, Δημήτριος Κουτσομητρόπουλος

από την προηγούμενη γενιά, όσο και εντελώς καινούρια κομμάτια που δοκιμάζονται για τυχόν καλή απόδοσή

τους. Με αυτόν τον τρόπο αξιοποιούν τις πληροφορίες που τους παρέχει το περιβάλλον τους.

Επανειλημμένες δοκιμές και πειράματα έχουν δείξει ότι μια «φυσική» αναπαράσταση των πιθανών

λύσεων για ένα δεδομένο πρόβλημα σε συνδυασμό με την εφαρμογή σε αυτή μιας οικογένειας γενετικών

τελεστών αποτελεί εξαιρετικά χρήσιμο εργαλείο στην προσπάθεια προσέγγισης των πραγματικών λύσεων σε

μια πολύ μεγάλη ποικιλία προβλημάτων και εφαρμογών. Αυτό το γεγονός μετατρέπει την προσέγγιση «φυσικού

μοντέλου» σε μια πολλά υποσχόμενη κατεύθυνση όσον αφορά την επίλυση προβλημάτων γενικότερα.

(ΘΕΩΡΙΑ ΕΞΕΛΙΞΓΗΣ ΤΩΝ ΕΙΔΩΝ - ΛΥΚΟΘΑΝΑΣΗΣ)

Η θεωρία της *Εξέλιξης των Ειδών (Evolution of Species)* που αναπτύχθηκε από τον Δαρβίνο στα μέσα του

περασμένου αιώνα προκάλεσε μεγάλη αναστάτωση, αφού ερχόταν σε σύγκρουση με τις επικρατούσες

θρησκευτικές αντιλήψεις περί προέλευσης της ζωής. Με την πάροδο ενός και πλέον αιώνα ο θόρυβος αυτός

δεν έχει κοπάσει πλήρως, όμως η θεωρία έχει γίνει αποδεκτή από το σύνολο των επιστημόνων, γιατί κατόρθωσε

να πείσει και να δώσει ικανοποιητικές απαντήσεις σε θεμελιώδη ερωτήματα. Σκοπός της είναι να δώσει μια

εξήγηση για το φαινόμενο της ζωής, την προέλευσή της και τις βασικές λειτουργίες της. Τα κυριότερά της

σημεία που σχετίζονται με τον τρόπο λειτουργίας των Γενετικών Αλγορίθμων και τον ερμηνεύουν είναι τα

εξής:

Δεν υπάρχει αντικειμενική βάση διαχωρισμού των ζωντανών οργανισμών σε ανώτερους και κατώτερους.

Σε κάθε βιολογικό είδος μερικά άτομα αφήνουν περισσότερους απογόνους σε σύγκριση με τα υπόλοιπα και

έτσι τα κληροδοτούμενα χαρακτηριστικά των αναπαραγωγικά επιτυχημένων ατόμων γίνονται περισσότερα

στην επόμενη γενιά. Οι δυσκολίες, τα εμπόδια και οι αντιξοότητες που παρουσιάζονται κατά τη διάρκεια της

ζωής των οργανισμών είναι οι παράγοντες που καθορίζουν ποιοι από αυτούς θα κατορθώσουν να ζήσουν και

να πολλαπλασιαστούν. Έτσι με την αλλαγή του περιβάλλοντος και των συνθηκών διαβίωσης αλλάζουν και τα

χαρακτηριστικά τους, προσπαθώντας να προσαρμοστούν κάθε φορά, με στόχο την εξασφάλιση της επιβίωσής

τους.

Αυτή η αλλαγή όμως που συμβαίνει στηα χαρακτηριστικά των ατόμων είναι αλλαγή στα *χρωμοσώματά*

τους (*chromosomes*), τα πολύπλοκα οργανικά μόρια που κωδικοποιούν τη δομή και τα χαρακτηριστικά τους.

Τα χρωμοσώματα αποτελούνται από μικρότερα μέρη, γνωστά ως *γονίδια (genes)*. Το σύνολο της γενετικής

πληροφορίας που είναι κωδικοποιημένο στα γονίδια ονομάζεται *γονότυπος ή γενότυπος (genotype)*. Η

δημιουργία ενός νέου οργανισμού περιλαμβάνει την αποκωδικοποίηση των χρωμοσωμάτων. Το σύνολο των

«ορατών» χαρακτηριστικών του και της συμπεριφοράς του, που καθορίζονται από τις πληροφορίες των

γονιδίων, συνιστούν τον *φαινότυπο (phenotype)*.

Κυρίαρχες λειτουργίες του φαινομένου της εξέλιξης είναι η *αναπαραγωγή (reproduction)* και η

*μετάλλαξη (mutation)*. Κατά τη μετάλλαξη γίνεται με τυχαίο τρόπο η αλλαγή της δομής των χρωμοσωμάτων,

συνήθως από λανθασμένη αντιγραφή βιολογικών μορίων ή από εξωγενείς παράγοντες (π.χ. ακτινοβολία),

έχοντας ως άμεσο αποτέλεσμα αλλαγή σε κάποιο χαρακτηριστικό. Η μετάλλαξη μερικές φορές μπορεί να

προκαλέσει βελτιώσεις και, χωρίς αμφιβολία, μερικά λάθη που έγιναν αποτέλεσαν σημαντικό παράγοντα για

την προοδευτική εξέλιξη της ζωής.

Προϊόν της αναπαραγωγής είναι ένας νέος οργανισμός, τα χρωμοσώματα του οποίου αποτελούνται από

γονίδια που προέρχονται τα μισά από τον πατέρα και τα μισά από τη μητέρα. Έτσι για κάθε χαρακτηριστικό το

νέο άτομο έχει πάρει ένα γονίδιο από κάθε γονέα. Μερικές φορές τα δύο αυτά γονίδια συμφωνούν μεταξύ τους

όσον αφορά την «τιμή» που θα δώσουν στο χαρακτηριστικό, π.χ. γαλάζιο χρώμα ματιών, ενώ άλλες φορές δεν

συμφωνούν, π.χ. το ένα υποδεικνύει καστανό χρώμα ματιών και το άλλο γαλάζιο. Στη δεύτερη περίπτωση

κυριαρχεί η «τιμή» ενός γονιδίου (π.χ. του καστανού) και αγνοείται η «τιμή» του άλλου, μολονότι το δεύτερο

μπορεί να περάσει σε επόμενες γενιές. Το γονίδιο που τελικά καθορίζει το χαρακτηριστικό λέγεται *κυρίαρχο* ή

*επικρατές (dominant)* και το άλλο *υπολειπόμενο (recessive)*. Γονίδια που διεκδικούν την ίδια θέση σε ένα

χρωμόσωμα (δηλαδή που είναι υπεύθυνα για το ίδιο χαρακτηριστικό) λέγονται *αλληλόμορφα (alleles).*

Όλος αυτός ο μηχανισμός της φυσικής επιλογής φάνηκε ιδιαίτερα ελκυστικός στον John Holland,

πρωτοπόρο των Γενετικών Αλγορίθμων, στις αρχές της δεκαετίας του ’70 (Goldberg, 1989). Ο Holland

φαντάστηκε ότι κάποιες ιδέες και λειτουργίες που εφαρμόζει η φύση στα συστήματά της θα μπορούσαν να

έχουν αποτελέσματα, αν ενσωματώνονταν σε αλγόριθμους για υπολογιστές, ώστε να προκύψουν αποδοτικές

τεχνικές επίλυσης δύσκολων προβλημάτων. Αποτέλεσμα αυτής της εργασίας του Ηolland ήταν οι Γενετικοί

Αλγόριθμοι, μια καινούρια εξελισσόμενη και πολλά υποσχόμενη τεχνική αναζήτησης και βελτιστοποίησης.

Η βασική ιδέα που κρύβεται πίσω από τους Γενετικούς Αλγόριθμους (ΓΑ) είναι η μίμηση των

μηχανισμών της φύσης. Ας πάρουμε π.χ. τους λαγούς και πώς αναπαράγονται και εξελίσσονται από γενιά σε

γενιά. Έστω ότι αρχίζουμε να παρατηρούμε έναν συγκεκριμένο πληθυσμό από λαγούς. Όπως είναι φυσικό,

κάποιοι από αυτούς θα είναι πιο γρήγοροι και πιο εύστροφοι από τους άλλους. Αυτοί οι γρηγορότεροι και

εξυπνότεροι λαγοί έχουν λιγότερες πιθανότητες να αποτελέσουν γεύμα κάποιας αλεπούς, άρα από τη στιγμή

που καταφέρνουν να επιβιώσουν θα ασχοληθούν με την αναπαραγωγή του είδους τους. Φυσικά θα υπάρχει και

ένας μικρός αριθμός αργών και λιγότερο εύστροφων λαγών που θα καταφέρουν να επιβιώσουν μόνο και μόνο

επειδή στάθηκαν τυχεροί. Όλοι αυτοί οι λαγοί που έχουν καταφέρει να επιβιώσουν θα αρχίσουν την παραγωγή

της επόμενης γενιάς τους, μιας γενιάς που θα συνδυάζει με διάφορους τρόπους όλα τα χαρακτηριστικά των Υπολογιστική νοημοσύνη και βαθιά μάθηση

25

μελών της προηγούμενης. Έτσι μερικοί αργοί λαγοί θα αναμειχθούν με κάποιους γρήγορους, κάποιοι γρήγοροι

με άλλους γρήγορους, κάποιοι εύστροφοι λαγοί με κάποιους μη εύστροφους και ούτω καθεξής. Οι μικροί λαγοί

της επόμενης γενιάς θα είναι κατά μέσο όρο γρηγορότεροι και εξυπνότεροι από τους προγόνους τους, αφού

από την προηγούμενη γενιά επιβίωσαν περισσότεροι γρήγοροι και έξυπνοι λαγοί. Ευτυχώς για τη διατήρηση

της φυσικής ισορροπίας, και οι αλεπούδες υφίστανται την ίδια διαδικασία αναπαραγωγής, διαφορετικά οι λαγοί

θα γινόντουσαν υπερβολικά γρήγοροι και έξυπνοι για να μπορούν να τους πιάσουν.

## Ιστορική Αναδρομή

## 3.2 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]

## 3.3 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]

## 3.4 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]

## 3.5 Πλεονεκτήματα γενετικών αλγορίθμων

(Λυκοθανάσης)

Η χρήση των ΓΑ σε διάφορες εφαρμογές είναι ελκυστική για αρκετούς λόγους. Οι κυριότεροι ίσως είναι οι

εξής:

1) **Μπορούν να λύσουν δύσκολα προβλήματα γρήγορα και αξιόπιστα**. Ένας από τους σημαντικούς λόγους

χρήσης των ΓΑ είναι η μεγάλη τους αποδοτικότητα. Τόσο η θεωρία όσο και η πράξη έχουν δείξει ότι

προβλήματα που έχουν πολλές, δύσκολα προσδιορισμένες λύσεις μπορούν να αντιμετωπιστούν καλύτερα

από ΓΑ. Είναι δε αξιοσημείωτο ότι συναρτήσεις που παρουσιάζουν μεγάλες διακυμάνσεις και καθιστούν

ανεπαρκείς άλλες μεθόδους στην εύρεση των ακρότατών τους, για τους ΓΑ αυτές οι διακυμάνσεις δεν

αποτελούν σημεία δυσχέρειας.

2) **Μπορούν εύκολα να συνεργαστούν με τα υπάρχοντα μοντέλα και συστήματα.** Οι ΓΑ προσφέρουν το

σημαντικό πλεονέκτημα της χρήσης τους με προσθετικό τρόπο στα μοντέλα που χρησιμοποιούνται σήμερα,

μην απαιτώντας την επανασχεδίασή τους. Μπορούν εύκολα να συνεργαστούν με τον υπάρχοντα κώδικα,

χωρίς μεγάλο κόπο. Αυτό συμβαίνει διότι χρησιμοποιούν μόνο πληροφορίες της διαδικασίας ή συνάρτησης

που πρόκειται να βελτιστοποιήσουν, δίχως να ενδιαφέρει άμεσα ο ρόλος της μέσα στο σύστημα ή η όλη

δομή του συστήματος.

3) **Είναι εύκολα επεκτάσιμοι και εξελίξιμοι.** Όπως θα γίνει σαφές στα επόμενα κεφάλαια, οι ΓΑ δεν

αντιστέκονται σε αλλαγές, επεκτάσεις και μετεξελίξεις, ανάλογα με την κρίση του σχεδιαστή. Σε πολλές

εφαρμογές έχουν αναφερθεί λειτουργίες των ΓΑ που δεν είναι δανεισμένες από τη φύση ή που έχουν υποστεί

σημαντικές αλλαγές, πάντα προς όφελος της απόδοσης. Παραλλαγές στο βασικό σχήμα δεν είναι απλά

ανεκτές, αλλά σε ορισμένες περιπτώσεις επιβάλλονται.

4) **Μπορούν να συμμετέχουν σε υβριδικές μορφές με άλλες μεθόδους**. Αν και η ισχύς των ΓΑ 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 ΓΑ με

άλλη μέθοδο. Αυτό είναι αποτέλεσμα της μεγάλης ευελιξίας των ΓΑ.

5) **Εφαρμόζονται σε πολύ περισσότερα πεδία από κάθε άλλη μέθοδο.** Το χαρακτηριστικό που τους

εξασφαλίζει αυτό το πλεονέκτημα είναι η ελευθερία επιλογής των κριτηρίων που καθορίζουν την επιλογή

μέσα στο τεχνικό περιβάλλον. Έτσι ΓΑ μπορούν να χρησιμοποιηθούν στην οικονομία, στον σχεδιασμό

μηχανών, στην επίλυση μαθηματικών εξισώσεων, στην εκπαίδευση Νευρωνικών Δικτύων και σε πολλούς

άλλους τομείς.

6) **Δεν απαιτούν περιορισμούς στις συναρτήσεις που επεξεργάζονται**. Ο κύριος λόγος που καθιστά τις

παραδοσιακές μεθόδους δύσκαμπτες και ακατάλληλες για πολλά προβλήματα είναι η απαίτησή τους για

ύπαρξη περιορισμών, όπως η ύπαρξη παραγώγων, η συνέχεια, οι μη «θορυβώδεις» συναρτήσεις κλπ.

Τέτοιου είδους ιδιότητες είναι αδιάφορες για τους ΓΑ, στοιχείο που τους κάνει κατάλληλους για μεγάλο

φάσμα προβλημάτων.

7) **Δεν ενδιαφέρει η σημασία της υπό εξέταση πληροφορίας.** Η μόνη «επικοινωνία» του ΓΑ με το

περιβάλλον του είναι η αντικειμενική συνάρτηση. Αυτό εγγυάται την επιτυχία του ανεξάρτητα από τη

σημασία του προβλήματος. Βέβαια δεν σημαίνει ότι δεν υπάρχουν άλυτα προβλήματα για τους ΓΑ. Όπου

όμως δεν τα καταφέρνουν, η αιτία είναι η φύση του χώρου που ερευνούν και όχι το πληροφοριακό

περιεχόμενο του προβλήματος.

8) **Έχουν από τη φύση τους το στοιχείο του παραλληλισμού.** Οι ΓΑ σε κάθε τους βήμα επεξεργάζονται

μεγάλες ποσότητες πληροφορίας, αφού κάθε άτομο θεωρείται αντιπρόσωπος πολλών άλλων. Έχει

υπολογιστεί ότι η αναλογία αυτή είναι της τάξεως *O*(*n* 3 ), δηλαδή 10 άτομα αντιπροσωπεύουν περίπου 1000.

Επίσης, είναι η μοναδική μέθοδος που εξετάζει παράλληλα πολλά σημεία (όσα και τα άτομα του 26

Σπυρίδων Λυκοθανάσης, Δημήτριος Κουτσομητρόπουλος

πληθυσμού). Είναι λοιπόν προφανές ότι μπορούν να καλύψουν με αποδοτικό ψάξιμο μεγάλους χώρους σε

μικρούς χρόνους.

9) **Είναι η μόνη μέθοδος που κάνει ταυτόχρονα εξερεύνηση του χώρου αναζήτησης και εκμετάλλευση**

**της ήδη επεξεργασμένης πληροφορίας.** Ο συνδυασμός αυτός σπάνια συναντάται σε οποιαδήποτε άλλη

μέθοδο. Με το τυχαίο ψάξιμο γίνεται καλή εξερεύνηση του χώρου, αλλά δεν γίνεται εκμετάλλευση της

πληροφορίας. Αντίθετα, με την αναζήτηση με μικρά άλματα στη συνάρτηση (hillclimbing) γίνεται καλή

εκμετάλλευση της πληροφορίας, αλλά όχι καλή εξερεύνηση. Συνήθως τα δύο αυτά χαρακτηριστικά είναι

ανταγωνιστικά και το επιθυμητό είναι να συνυπάρχουν και τα δύο προς όφελος της όλης διαδικασίας. Οι

ΓΑ επιτυγχάνουν τον βέλτιστο συνδυασμό εξερεύνησης και εκμετάλλευσης, κάτι που τους κάνει ιδιαίτερα

αποδοτικούς και ελκυστικούς.

10) **Επιδέχονται παράλληλη υλοποίηση.** Οι ΓΑ μπορούν να εκμεταλλευτούν τα πλεονεκτήματα των

παράλληλων μηχανών, αφού λόγω της φύσης τους εύκολα μπορούν να δεχτούν παράλληλη υλοποίηση. Το

χαρακτηριστικό αυτό αυξάνει ακόμη περισσότερο την απόδοσή τους, ενώ σπάνια συναντάται σε

ανταγωνιστικές μεθόδους.

## 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. (Haykin, 2010) & (Θεοδωρίδης & Κουτρούμπας, 2012) & (Κίτρου, χ.χ.) & (Διαμαντάρας, 2007) & (Λυκοθανάσης, 2021) [↑](#footnote-ref-7)
9. (Τσούλος, χ.χ.) & (Βλαχάβας et al., 2006) & (Ντούνης, 2024) [↑](#footnote-ref-8)
10. (Haykin, 2010) [↑](#footnote-ref-9)
11. (Rosenblatt, 1958) [↑](#footnote-ref-10)
12. (Διαμαντάρας, 2007) [↑](#footnote-ref-11)
13. (Βλαχάβας et al., 2006) [↑](#footnote-ref-12)
14. (Διαμαντάρας, 2007) [↑](#footnote-ref-13)
15. (Διαμαντάρας, 2007) [↑](#footnote-ref-14)
16. (Haykin, 2010) [↑](#footnote-ref-15)