

**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

**TRAINING OF ARTIFICIAL NEURAL NETWORKS WITH STOCHASTIC TECHNIQUES**

**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

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 vi](#_Toc7197)

[ΠΕΡΙΛΗΨΗ vii](#_Toc20000)

[ABSTRACT viii](#_Toc14790)

[ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ ix](#_Toc10738)

[ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ xi](#_Toc22403)

[ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ xii](#_Toc20039)

[ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ xiii](#_Toc16002)

[1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα 1](#_Toc14186)

[1.1 Ιστορική Αναδρομή 1](#_Toc23129)

[1.2 Μετάβαση από τους βιολογικούς στους τεχνητούς νευρώνες 2](#_Toc18756)

[1.3 Συναρτήσεις ενεργοποίησης 5](#_Toc13135)

[1.4 Τεχνητά Νευρωνικά Δίκτυα: Μία σφαιρική εικόνα 8](#_Toc22266)

[1.5 Μάθηση με επίβλεψη 11](#_Toc29701)

[2. Δομή και Λειτουργία των Τεχνητών Νευρωνικών Δικτύων 13](#_Toc15476)

[2.1 Τα δίκτυα Perceptron 13](#_Toc7522)

[2.2 Τα δίκτυα Adaline 16](#_Toc16663)

[2.3 Τα δίκτυα MLP 19](#_Toc25327)

[2.4 Η μέθοδος Back Propagation 23](#_Toc29998)

[2.5 Η μέθοδος Gradient Descent 26](#_Toc15113)

[2.6 Εφαρμογές νευρωνικών δικτύων 27](#_Toc2018)

[3. Γενετικοί αλγόριθμοι 29](#_Toc18824)

[3.1 Ιστορική Αναδρομή 29](#_Toc16817)

[3.2 Μέθοδοι κωδικοποίησης 30](#_Toc16599)

[3.3 Συνάρτηση καταλληλότητας 30](#_Toc28600)

[3.4 Γενετικοί τελεστές 31](#_Toc29836)

[3.5 Εφαρμογές γενετικών αλγορίθμων 33](#_Toc11866)

[4. Μέθοδος - αποτελέσματα 36](#_Toc29185)

[4.1 Τα dataset που χρησιμοποιήθηκαν 36](#_Toc13828)

[4.2 Η προτεινόμενη μέθοδος 37](#_Toc22148)

[4.3 Πειραματικά αποτελέσματα 38](#_Toc8987)

[5. Συμπεράσματα 39](#_Toc6152)

[ΒΙΒΛΙΟΓΡΑΦΙΑ 40](#_Toc14936)

[ΠΑΡΑΡΤΗΜΑ 42](#_Toc23694)

# ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1.1. Βιολογικός νευρώνας…………………..….…………...….............03

Εικόνα 1.2. Τεχνητός νευρώνας...……………………………......….................05

Εικόνα 1.3. Γραφική παράσταση της γραμμικής συνάρτησης………................06

Εικόνα 1.4. Γραφική παράσταση της βηματικής συνάρτησης 0/1.…………….06

Εικόνα 1.5. Γραφική παράσταση της βηματικής συνάρτησης -1/1….……........07

Εικόνα 1.6. Γραφική παράσταση της συνάρτησης σιγμοειδής συνάρτησης.......07

Εικόνα 1.7. Γραφική παράσταση της συνάρτησης υπερβολικής εφαπτομένης..08

Εικόνα 1.6. Σχηματική απεικόνιση ενός απλού ΤΝΔ..........................................09

Εικόνα 2.1. Perceptron..........................................……………….......................16

Εικόνα 2.2. MLP……...............................................................………………...16

# ΠΙΝΑΚΑΣ ΣΥΝΤΟΜΟΓΡΑΦΙΩΝ

**ΤΝΔ**............................................................................Τεχνητά Νευρωνικά Δίκτυα

**MLP**......................................................Μοντέλο Perceptron πολλών στρωμάτων

**LMS**.........................................................................................Least Mean Squares

# ΑΠΟΔΟΣΗ ΟΡΩΝ / ΓΛΩΣΣΑΡΙΟ

**On-Line Αλγόριθμος Μάθησης:** Η εκπαίδεση των βαρών γίνεται με κάθε πέρασμα ενός προτύπου και όχι μετά την ολοκλήρωση μίας ολόκληρης εποχής.

**ΠΡΟΤΥΠΑ:** Το σύνολο των δεδομένων εισόδου.

**ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΠΡΟΤΥΠΟΥ**: Κάθε τιμή που αναπαριστά μία ιδιότητα.

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

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

**ΚΛΕΙΣΤΟ ΔΙΑΣΤΗΜΑ**: Το κλειστό διάστημα [α,β] μίας μεταβλητής x σημαίνει πως wps.

# 

# 1. Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα

## Ιστορική Αναδρομή

Τη δεκαετία του 1940 σημειώθηκε σημαντική πρόοδος στη μελέτη των βιολογικών νευρωνικών δικτύων και στη μαθηματική μοντελοποίησή τους. Η περίοδος των ΤΝΔ ξεκίνησε το 1943 με την ανάπτυξη ενός βασικού μοντέλου τεχνητού νευρώνα, από τους Αμερικανούς επιστήμονες McCulloch και Pitts. 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 στοιχεία αυτού του τεχνητού νευρώνα για την κατασκευή του EDVAC, του πρώτου γενικού σκοπού υπολογιστή το 1949.

Οκτώ χρόνια αργότερα, το 1957, ο ψυχολόγος Frank Rosenblatt ανέπτυξε το δίκτυο Perceptron στο Αεροναυτικό Εργαστήριο του Κορνέλλ (Cornell Aeronautical Laboratory). Αποτελεί μία καινοτόμα μέθοδο μάθησης με επίβλεψη και είναι μία ελαφρώς τροποποιημένη έκδοση του νευρώνα των McCulloch και Pitts. Το δίκτυο αυτό, δημοσιεύτηκε το 1958 και 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. Αυτόν τον περιορισμό ήρθε να καταργήσει το δίκτυο Adaline. Αναπτύχθηκε το 1960, από τους Widrow και Hoff και αποτελεί μία τροποποίηση του δικτύου Perceptron. Δίκτυο, που για πρώτη φορά μπορεί να εφαρμοστεί σε μη γραμμικά διαχωρίσιμα πρότυπα με επιτυχία. Έτσι, αποτέλεσε ένα από τα πρώτα δίκτυα που εφαρμόστηκαν σε πρακτικά προβλήματα. Ειδικότερα, χρησιμοποιήθηκε ως φίλτρο για την εξάλειψη της ηχούς σε τηλεφωνικές γραμμές καθώς και για την πρόβλεψη και την ισοστάθμιση καναλιών επικοινωνίας.

Τη δεκαετία του 1970, η έρευνα πάνω στα ΤΝΔ παρουσίασε ύφεση. Όμως, κατά τη δεκαετία του 1980, σημειώθηκε έκρηξη δραστηριότητας με την εισαγωγή περισσότερων νευρώνων σε ένα δίκτυο Perceptron. Το νέο δίκτυο που προέκυψε ονομάστηκε MLP και προσφέρει ακόμη και σήμερα απεριόριστες δυνατότητες αναπαράστασης συναρτήσεων και διαχωρισμού κλάσεων ξεπερνώντας τους υπολογιστικούς περιορισμούς που είχαν παρουσιαστεί για το δίκτυο Perceptron.

Την ίδια περίοδο, ο Paul Werbos πρότεινε τον αλγόριθμο εκπαίδευσης Back-Propagation στα πλαίσια της ανάλυσης μοντέλων οικονομικής και πολιτικής πρόβλεψης. Τότε, τα μοντέλα αυτά δεν είχαν σχεδιαστεί για να έχουν σχέση ή να θυμίζουν ΤΝΔ. Αργότερα, μέσα στη δεκαετία του 1980, έγινε αντιληπτό ότι ο αλγόριθμος αυτός, μπορούσε να μεταφερθεί αυτούσιος στην εκπαίδευση των MLP δικτύων και έκτοτε έγινε ο πιο δημοφιλής και διαδεδομένος για το σκοπό αυτό και έδωσε νέα ώθηση στις εφαρμογές ΤΝΔ. Αλγόριθμος, ικανός για πρώτη φορά να εκπαιδεύσει ένα δίκτυο με περισσότερους από έναν νευρώνες. Η ανάπτυξή του, αποτέλεσε ορόσημο στην εξέλιξη του κλάδου των ΤΝΔ.

Τα ΤΝΔ 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 (Haykin, 2010)· (Γεωργούλη, 2015)· (Διαμαντάρας, 2007)· (Λυκοθανάσης και Κουτσημητρόπουλος, 2021)· (Ντούνης, 2024)· (Τσίπουρας, 2015)· (Τσούλος, χ.χ.)

## 1.2 Μετάβαση από τους βιολογικούς στους τεχνητούς νευρώνες

Ο ανθρώπινος εγκέφαλος αποτελεί το πιο πολύπλοκο όργανο του ανθρώπινου σώματος και είναι υπεύθυνος για τις πιο σύνθετες και εξελιγμένες λειτουργίες που αυτό εκτελεί όπως η μνήμη, η σκέψη, και η ανάλυση. Λαμβάνει πληροφορίες μέσω των αισθήσεων όπως είναι η όραση και η ακοή και εκπέμπει κινητικές διεγέρσεις για τον συντονισμό των κινήσεων καθώς και τη ρύθμιση των σωματικών λειτουργιών (Δούνιας και Καραμπότσης, χ.χ.) 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 τη 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 - ο οποίος έχει το σχήμα μακρόστενης κλωστής. Ανάμεσα στον άξονα του νευρώνα και στους δενδρίτες άλλων νευρώνων, υπάρχει ένα σημείο ένωσης, που ονομάζεται σύναψη. Οι συνάψεις αυτές - 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. Χαρακτηριστικά όπως το πλάτος της σύναψης, η απόστασή της από τον δενδρίτη επηρεάζουν την ευκολία με την οποία μεταδίδεται εν τέλει το φορτίο. Η ευκολία αυτή εκφράζεται με μία τιμή, την οποία αποκαλούμε 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 ανάμεσα τους, τόσο πιο έντονα συμμετέχει το συγκεκριμένο φορτίο εισόδου στο συνολικό άθροισμα. Αν το άθροισμα του φορτίου ξεπερνάει κάποιο κατώφλι (threshold) 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 αυτό, 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. (Βλ. Εικόνα 1.1 για τη δομή του βιολογικού νευρώνα) (Haykin, 2010)· (Βλαχάβας, Κεφάλας, Βασιλειάδης, Κόκκορας και Σακελλαρίου, 2006)· (Διαμαντάρας, 2007)



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

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 ικανότητά 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. (Haykin, 2010)· (Βλαχάβας et al., 2006)· (Πούλος, 2015)

Αξίζει να σημειωθεί πως τα συνήθη ΤΝΔ χρησιμοποιούν απλοποιημένα μοντέλα νευρώνων, τα οποία διατηρούν μόνο τα πολύ βασικά χαρακτηριστικά των λεπτομερών μοντέλων που χρησιμοποιούνται στη νευρολογία. Θα έλεγε κανείς πως οι τεχνητοί νευρώνες που χρησιμοποιούμε για την κατασκευή των ΤΝΔ είναι αρκετά πρωτόγονοι σε σύγκριση με τους νευρώνες του ανθρώπινου εγκεφάλου. Ωστόσο, υπάρχει η πεποίθηση πως παρά την απλότητα των ΤΝΔ, οι λεπτομέρειες δεν έχουν ιδιαίτερη σημασία στην κατανόηση της ευφυούς συμπεριφοράς των βιολογικών νευρωνικών συστημάτων. Αυτά τα απλά μοντέλα νευρώνων αξιοποιώντας δύο (2) βασικά χαρακτηριστικά, μπορούν να δημιουργήσουν ιδιαιτέρως ενδιαφέροντα δίκτυα. Το πρώτο χαρακτηριστικό είναι η διευκόλυνση της μάθησης με τη χρήση ρυθμιζόμενων παραμέτρων στους νευρώνες, το οποίο είναι γνωστό και ως πλαστικότητα των νευρώνων. Το δεύτερο χαρακτηριστικό, είναι πως το δίκτυο αποτελείται από μεγάλο πλήθος νευρώνων ώστε να επιτυγχάνεται παραλληλισμός της επεξεργασίας και κατανομή της πληροφορίας. (Διαμαντάρας, 2007)· (Ηλιάδης, χ.χ.)

Ο τεχνητός νευρώνας λοιπόν, είναι ένα υπολογιστικό μοντέλο, τα μέρη του οποίου μπορούν να αντιστοιχιστούν άμεσα με αυτά του βιολογικού νευρώνα. Δέχεται αντί για ηλεκτρικούς παλμούς, συνεχείς μεταβλητές ως σήματα εισόδου. Κάθε σήμα εισόδου, μεταβάλλεται από μία τιμή βάρους η οποία είναι πραγματικός αριθμός, και αντιστοιχεί στον 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. Συνήθως 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. Μπορούμε να πούμε πως ο νευρώνας είναι ντετερμινιστικό (αιτιοκρατικό) μοντέλο, διότι είναι επακριβώς καθορισμένη η συμπεριφορά του για όλες τις εισόδους. Δηλαδή για κάθε είσοδο, υπάρχει μία προκαθορισμένη έξοδος. (Βλ. Εικόνα 1.2 για τη δομή του τεχνητού νευρώνα) (Haykin, 2010)· (Βλαχάβας et al., 2006)· (Πούλος, 2015)



Εικόνα 1.2: Αναπαράσταση τεχνητού νευρώνα (Suzuki, 2011)

## 1.3 Συναρτήσεις ενεργοποίησης

Υπάρχουν αρκετές προσεγγίσεις για τη μοντελοποίηση του νευρώνα, οι οποίες διαφέρουν από το μοντέλο McCulloch και Pitts. Η πιο σημαντική διαφορά είναι στη μορφή της συνάρτησης ενεργοποίησης που χρησιμοποιείται. Όπως αναφέρθηκε στην προηγούμενη ενότητα, η συνάρτηση, η οποία συμβολίζεται ως f(x), αποτελεί ένα από τα δύο βασικά τμήματα του τεχνητού νευρώνα, καθώς ορίζει την έξοδο του. Υπάρχουν αρκετά είδη συναρτήσεων που μπορούν να χρησιμοποιηθούν, με πιο δημοφιλή τα παρακάτω:



Γραμμική συνάρτηση:[[1]](#footnote-0)

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

Figure 3

(Εικόνα 1.3) Γραφική παράσταση της γραμμικής συνάρτησης



Βηματική συνάρτηση:[[2]](#footnote-1)



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



Figure 4

Αλλά και με την μορφή [[3]](#footnote-2)

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



Figure 5

Σιγμοειδής συνάρτηση:[[4]](#footnote-3)



(Εικόνα 1.6) Γραφική παράσταση της σιγμοειδής συνάρτησης



Figure 6

Υπερβολική εφαπτομένη:



(Εικόνα 1.7) Γραφική παράσταση της συνάρτησης υπερβολικής εφαπτομένης

## 1.4 Τεχνητά Νευρωνικά Δίκτυα: Μία σφαιρική εικόνα

Τα ΤΝΔ μοιράζονται δύο (2) χαρακτηριστικά με τον ανθρώπινο εγκέφαλο. Το πρώτο από αυτά είναι πως το δίκτυο προσλαμβάνει τη γνώση από το περιβάλλον του, μέσω μίας διαδικασίας μάθησης. Το δεύτερο χαρακτηριστικό είναι πως η αποκτηθείσα γνώση αποθηκεύεται μέσω της προσαρμογής της ισχύος των συνδέσεων μεταξύ των νευρώνων. Χάριν συντομίας, τα δίκτυα αυτά σύμφωνα με τον Haykin, αποκαλούνται και νευρωνικά δίκτυα. Σύμφωνα με τον Βλαχάβα, έχει διατυπωθεί ο παρακάτω ορισμός από τους Aleksander και Morton:

*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.*

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

Υπάρχουν δύο κατηγορίες νευρώνων ανάλογα τον τρόπο διασύνδεσης τους στο δίκτυο. Η πιο συνηθισμένη περίπτωση είναι αυτή των πλήρως συνδεδεμένων νευρώνων κατά την οποία οι νευρώνες ενός στρώματος είναι πλήρως συνδεδεμένοι με τους νευρώνες του επόμενου στρώματος χωρίς όμως να είναι συνδεδεμένοι με νευρώνες είτε του προηγούμενου είτε του ίδιου στρώματος. Χαρακτηριστικό παράδειγμα αυτής της κατηγορίας, είναι τα δίκτυα πρόσθιας τροφοδότησης (feedforward) τα οποία χρησιμοποιούνται στην πλειοψηφία των εφαρμογών των ΤΝΔ. Το όνομα τους οφείλεται στο ότι η ροή της πληροφορίας μέσα στο δίκτυο είναι μονής κατεύθυνσης. Σε αυτά τα δίκτυα υπάρχουν τα στρώματα εισόδου και εξόδου και προαιρετικά ένα ή περισσότερα κρυφά στρώματα. (Βλαχάβας et al., 2006) Για τις περιπτώσεις που δεν υπάρχουν κρυφά στρώματα, πολλές φορές χρησιμοποιείται για αυτά, ο όρος Perceptron. (Γεωργούλη, 2015) Σε οποιαδήποτε άλλη περίπτωση θεωρούμε ότι οι νευρώνες είναι μερικώς συνδεδεμένοι. Στο παρακάτω σχήμα, παρουσιάζεται μία απλή μορφή ΤΝΔ πρόσθιας τροφοδότησης. Στο στρώμα εισόδου υπάρχουν δύο νευρώνες (x1 και x2), στο κρυφό στρώμα τρεις νευρώνες (S1, S2, S3) και στο στρώμα εξόδου ένας νευρώνας (y). Σε αυτό το σημείο καλό θα ήταν να διευκρινίσουμε πως παρόλο που ο νευρώνας y έχει δύο εξόδους (y1 και y2), αυτές θα πρέπει να έχουν τις ίδιες ακριβώς τιμές.



(Εικόνα 1.8) Σχηματική απεικόνιση ενός απλού ΤΝΔ.

Υπάρχουν τέσσερις (4) ιδιότητες που συνδέονται άρρηκτα με τα ΤΝΔ. Η πρώτη από αυτές είναι η ικανότητά τους να μαθαίνουν μέσω παραδειγμάτων (learn by example). Η δεύτερη ιδιότητα αφορά τη δυνατότητα θεώρησής τους ως κατανεμημένη μνήμη (distributed memory) και ως μνήμη συσχέτισης (associative memory). Αυτό οφείλεται στο γεγονός ότι η κωδικοποίηση που δημιουργείται είναι κατανεμημένη σε όλα τα βάρη της συνδεσμολογίας τους. Επίσης, αντί για τον παραδοσιακό τρόπο αποθήκευσης της πληροφορίας, χρησιμοποιούν κατάλληλες συσχετίσεις που προκύπτουν από τα δεδομένα εκπαίδευσης. Έτσι, η ανάκληση της πληροφορίας γίνεται βάσει του περιεχομένου και όχι της διεύθυνσης, όπως ακριβώς συμβαίνει και 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 ελλιπή. Επιπλέον, ακόμη και σε περίπτωση δομικών σφαλμάτων, η κακή λειτουργία ή ακόμη και η καταστροφή ενός νευρώνα ή κάποιων συνδέσεων δε μπορεί να διαταράξει σημαντικά τη λειτουργία του δικτύου. Αυτό συμβαίνει καθώς η πληροφορία που εσωκλείουν δεν είναι συγκεντρωμένη σε συγκεκριμένο σημείο αλλά είναι διάχυτη σε όλο το δίκτυο. Η τελευταία ιδιότητα είναι η ικανότητά 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 δύο τελευταίες 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. (Βλαχάβας et al., 2006)· (Λύκας, χ.χ)

## 1.5 Μάθηση με επίβλεψη

H πρόκληση που αντιμετωπίζουν τα ΤΝΔ είναι η εύρεση κατάλληλων αλγορίθμων με σκοπό την εκπαίδευση τους καθώς και την ανάκληση της πληροφορίας που περιέχεται. Αυτό έχει ως στόχο την 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 τα ΤΝΔ στοχεύουν επιπλέον, στην αυτοπροσαρμογή σε νέες καταστάσεις και γνωστά περιβάλλοντα, καθώς και στη 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 και υποθέτουμε πως δεν υπάρχει κάποια εξωτερική "αυθεντία" η οποία γνωρίζει τις ορθές παραμέτρους για τη λύση του προβλήματος. Ωστόσο, μπορούμε να θεωρήσουμε ότι ο εκπαιδευτής έχει γνώση του περιβάλλοντος, η οποία αντιπροσωπεύεται από ένα δείγμα προτύπων, έχοντας αντιστοίχηση μεταξύ των εισόδων (ή αλλιώς ερεθισμάτων) - γνωστές και ως ανεξάρτητες μεταβλητές και της αντίστοιχης επιθυμητής απόκρισης (στόχος) - γνωστή και ως εξαρτημένη μεταβλητή. (Βλαχάβας et al., 2006) Αυτή η πληροφορία είναι γνωστή και ως a-priori (εκ των προτέρων) και τα πρότυπα αναφέρονται ως χαρακτηρισμένα (labeled). 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. (Haykin, 2010)· (Διαμαντάρας, 2007)· (Θεοδωρίδης και Κουτρούμπας, 2012)· (Κίτρου, χ.χ.)· (Λυκοθανάσης και Κουτσημητρόπουλος, 2021) 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. Πρέπει να επισημανθεί σε αυτό το σημείο πως όσο περισσότερα δεδομένα έχουμε στη διάθεσή 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 και 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. Τέλος, σύμφωνα με τον Κίτρου, η μάθηση με επίβλεψη χρησιμοποιείται και σε προβλήματα πρόγνωσης αλλά και διερμηνείας.

# Δομή και Λειτουργία των Τεχνητών Νευρωνικών Δικτύων

## Τα δίκτυα Perceptron

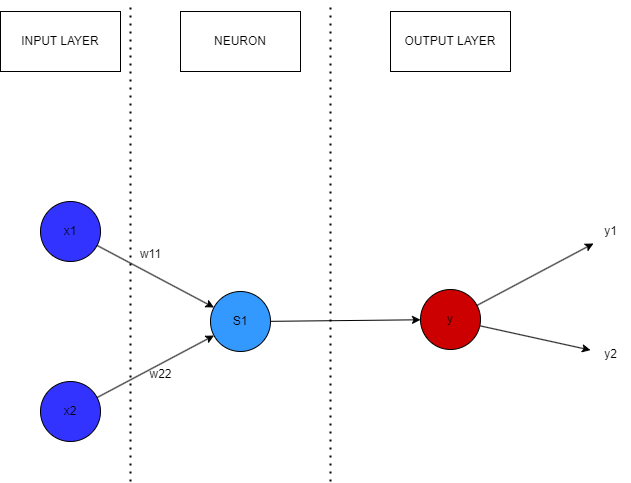
Τα δίκτυα Perceptron καταλαμβάνουν μία ειδική θέση στην ιστορική εξέλιξη των ΤΝΔ καθώς αποτελούν τη πρώτη μορφή δικτύου που μπορούσε να περιγραφεί αλγοριθμικά. Χρησιμοποιούνται για την ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, αρκεί αυτά να ανήκουν σε δύο μόνο κατηγορίες. Η σπουδαιότητα τους δεν είναι μόνο ιστορική καθώς συνεχίζουν να έχουν πρακτική χρησιμότητα. (Haykin, 2010) Υπάρχουν αρκετές παραλλαγές με την πιο απλή να είναι αυτή του στοιχειώδους Perceptron (elementary Perceptron). Κοινό χαρακτηριστικό σε όλες είναι πως περιλαμβάνουν μόνο έναν νευρώνα. Εφόσον δεν υπάρχουν παραπάνω από ένας νευρώνες ώστε να συνδεθούν μεταξύ τους, ο όρος δίκτυο χρησιμοποιείται καταχρηστικά.

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. Το τρίτο και τελευταίο ερώτημα, αφορά τον τρόπο με τον οποίο οι πληροφορίες που έχουν αποθηκευτεί επηρεάζουν την αναγνώριση και τη συμπεριφορά. Στόχος των δικτύων αυτών 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)

Έχει αποδειχθεί πως το αποτέλεσμα του Perceptron είναι το ίδιο, ανεξάρτητα από την συνάρτηση ενεργοποίησης που θα χρησιμοποιηθεί, αν και η πιο σύνηθες συνάρτηση για αυτό το δίκτυο είναι η βηματική. (Haykin, 2010)· (Βλαχάβας et al., 2006) Τα δίκτυα Perceptron αποτελούν την πιο απλή μορφή δικτύου πρόσθιας τροφοδότησης χρησιμοποιώντας μάθηση με επίβλεψη, η οποία σύμφωνα με τους Βλαχάβα et al είναι καθοδηγούμενη από το σφάλμα (error driven). 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). 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.

Βέβαια, αν το πρόβλημα δεν είναι γραμμικά διαχωρίσιμο τότε ο αλγόριθμος Perceptron δεν συγκλίνει ποτέ. Αυτό είναι ένα σοβαρό μειονέκτημα το οποίο αποτέλεσε και το βασικό σημείο κριτικής εναντίον του. Πολλά προβλήματα στον πραγματικό κόσμο είναι μη γραμμικά διαχωρίσιμα επομένως δεν μπορεί να χρησιμοποιηθεί αυτό το μοντέλο για αυτά. (Διαμαντάρας, 2007)

Στο παρακάτω σχήμα μπορούμε να δούμε μία επεικόνιση του δικτύου Perceptron με δύο (2) νευρώνες στο στρώμα εισόδου και έναν (1) νευρώνα στο στρώμα εξόδου. Καλό θα ήταν να σημειωθεί σε αυτό το σημείο πως συνηθίζεται η ύπαρξη ενός μόνο νευρώνα στο στρώμα εξόδου ενός δικτύου Perceptron, καθώς όπως είπαμε παραπάνω μπορεί να λειτουργήσει σωστά για την ταξινόμηση μόνο δύο (2) κατηγοριών.



(Εικόνα 2.1) Perceptron

Η υλοποίηση ενός δικτύου Perceptron, περιγράφεται παρακάτω με την μορφή ψευδοκώδικα:

ΕΙΣΟΔΟΙ:   
 Τα πρότυπα εισόδου με τις αντίστοιχες επιθυμητές εξόδους  
  
ΜΕΘΟΔΟΣ:

1. Αρχικοποίηση των βαρών που θα χρησιμοποιηθούν.
2. Αρχικοποίηση του ρυθμού μάθησης
3. Αρχικοποίηση του μέγιστου αριθμού επαναλήψεων - προαιρετικό
4. Όσο δεν ικανοποιείται η συνθήκη τερματισμού
   1. Για κάθε πρότυπο
      1. Υπολογισμός της εξόδου του δικτύου
      2. Αν (πραγματική έξοδος != επιθυμητή έξοδο)
         1. Εκπαίδευση των βαρών.

ΤΕΡΜΑΤΙΣΜΟΣ:  
 Αν δεν έχει χρειαστεί εκπαίδευση των βαρών για μία ολόκληρη εποχή ή

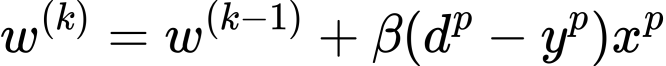
αν έχουμε ξεπεράσει τον μέγιστο αριθμό επαναλήψεων - εφόσον αυτός έχει οριστεί.

ΕΞΟΔΟΣ:

Τα ενημερωμένα βάρη

Αρχικά, πρέπει να αποθηκευτούν τα διανύσματα προτύπων που έχουμε στη διάθεση μας μαζί με τις αντίστοιχες επιθυμητές εξόδους. Αξίζει να σημειωθεί σε αυτό το σημείο, πως αν υπάρχει κάποιος λάθος στα πρότυπα ή στις επιθυμητές εξόδους τους, τότε αυτό θα οδηγήσει σε μη αξιόπιστα αποτελέσματα του δικτύου. Κάτι τέτοιο μπορεί να προκύψει από λανθασμένη εισαγωγή δεδομένων, ακόμη και από αστοχία υλικού - κάτι που μπορεί να γίνει για παράδειγμα από διαφόρων ειδών αισθητήρες. Κάποιες φορές μπορεί να υπάρχουν και ακραίες τιμές, το οποίο δυσκολεύει τη λειτουργία των ΤΝΔ. Το δεύτερο βήμα είναι η αρχικοποίηση του διανύσματος βαρών που θα χρησιμοποιεί ώστε να εκπαιδευτεί κατάλληλα προκειμένου να έχουμε το επιθυμητό αποτέλεσμα. Θα πρέπει πρώτα όμως να γνωρίζουμε το πλήθος των τιμών που χρειαζόμαστε για τα βάρη, το οποίο ισούται με το πλήθος των χαρακτηριστικών που έχουμε στη διάθεση μας αυξημένο κατά 1 για το bias. Υπάρχουν αρκετοί τρόποι αρχικοποίησης τους, με τον πιο απλό τρόπο αυτόν της αρχικοποίησης όλων των τιμών σε μία τιμή είτε ίση με το μηδέν (0) είτε με την μονάδα (1). Εναλλακτικά, μπορεί να γίνει αρχικοποίηση των τιμών σε τυχαίες τιμές στο κλειστό διάστημα [0,1]. Παράλληλα, γίνεται αρχικοποίηση του ρυθμού μάθησης του δικτύου σε μία μικρή θετική δεκαδική τιμή, συνήθως στο κλειστό διάστημα [0,1]. Συχνά προτιμούμε την τιμή 0.1. (Τσούλος, χ.χ.) Στη συνέχεια γίνεται μία επαναληπτική διαδικασία για κάθε πρότυπο που έχουμε στην διάθεση μας. Έτσι, για κάθε πρότυπο υπολογίζουμε την έξοδο του δικτύου χρησιμοποιώντας τις τιμές των βαρών που έχουμε στη διάθεση μας. Ο υπολογισμός αυτός προκύπτει από την συνάρτηση ενεργοποίησης που έχουμε επιλέξει ότι θα χρησιμοποιήσουμε στο δίκτυο. Συνήθως, χρησιμοποιείται η βηματική συνάρτηση. Ως είσοδος της συνάρτησης αυτής είναι ο υπολογισμός του αθροίσματος των εισόδων πολλαπλασιασμένων με τα αντίστοιχα συναπτικά βάρη. Σε περίπτωση που το αποτέλεσμα που έχουμε είναι το ίδιο με το επιθυμητό προχωράμε στο επόμενο πρότυπο. Αλλιώς, γίνεται εκπαίδευση των βαρών. Η διαδικασία αυτή τελειώνει είτε μετά από μία ολόκληρη εποχή που δεν έχει χρειαστεί εκπαίδευση των βαρών είτε μετά από κάποιο πεπερασμένο αριθμό επαναλήψεων - εφόσον έχει οριστεί στην αρχή της διαδικασίας.

Σχετικά με την ενημέρωση των βαρών όταν η πραγματική τιμή της εξόδου του δικτύου είναι διαφορετική από την επιθυμητή, γίνεται με την πρόσθεση ή την αφαίρεση ενός ποσοστού από το πρότυπο. Αυτό γίνεται με βάση τον παρακάτω μαθηματικό τύπο:

[[5]](#footnote-4)

## 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 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. Η χρήση όμως των συνεχών τιμών δημιουργεί μία επιπλέον πρόκληση. Για παράδειγμα, αν επιθυμούμε να χρησιμοποιήσουμε το δίκτυο Adaline για να διαχωρίσουμε δύο κατηγορίες, θα πρέπει να αποδώσουμε τις συνεχείς τιμές στις κατηγορίες, αν και συνήθως αυτό επιτυγχάνεται με τη διάκριση των θετικών και αρνητικών τιμών. Μοιάζουν αρκετά με τα δίκτυα Perceptron, αλλά μπορούν να χρησιμοποιηθούν επιπλέον για τη μάθηση συναρτήσεων καθώς και για τη μάθηση περισσότερων από δύο κατηγοριών υπό προυποθέσεις. Διαθέτουν και αυτά μόνο έναν νευρώνα. (Τσούλος, χ.χ.)

Τα δίκτυα Adaline χρησιμοποιούν τον αναδρομικό on-line αλγόριθμο εκπαίδευσης που είναι γνωστός ως LMS και προκύπτει από τη θεωρία στοχαστικών προσεγγίσεων. Υπάρχουν και άλλες ονομασίες για αυτόν, όπως κανόνας δέλτα (delta rule) ή κανόνας ADALINE ή ακόμη και κανόνας Widrow-Hoff από τα ονόματα αυτών που τον εισήγαγαν. Για πολύ καιρό αντιμετωπιζόταν ως μία “στιγμιαία” προσέγγιση της μεθόδου Gradient Descent που θα δούμε σε επόμενο υποκεφάλαιο. (Haykin, 2010) Η ελαχιστοποίηση σε αυτόν τον αλγόριθμο αφορά την ελαχιστοποίηση της τετραγωνικής απόστασης μεταξύ των διανυσμάτων της πραγματικής και επιθυμητής εξόδου.

Έχει παρατηρηθεί βέβαια, πως με την χρήση αυτού του δικτύου υπάρχει περίπτωση να μην μπορέσουμε να διαχωρίσουμε γραμμικά διαχωρίσιμα πρότυπα δύο (2) κατηγοριών, κάτι το οποίο κάνει με απόλυτη επιτυχία το δίκτυο Perceptron. Αλλά σε αντίθετη περίπτωση, το δίκτυο μπορεί να βρει ένα σχετικά καλό διάνυσμα βαρών ώστε να κατηγοριοποιήσει τα πρότυπα με αρκετά μεγάλο ποσοστό επιτυχίας. (Διαμαντάρας, 2007)

Η υλοποίηση ενός δικτύου Adaline, περιγράφεται παρακάτω με την μορφή ψευδοκώδικα:

ΕΙΣΟΔΟΙ:   
 Τα πρότυπα εισόδου με τις αντίστοιχες επιθυμητές εξόδους  
  
ΜΕΘΟΔΟΣ:

1. Αρχικοποίηση των βαρών που θα χρησιμοποιηθούν.
2. Αρχικοποίηση του ρυθμού μάθησης
3. Αρχικοποίηση κατωφλιού για το σφάλμα εκπαίδευσης
4. Αρχικοποίηση του μέγιστου αριθμού επαναλήψεων - προαιρετικό
5. Όσο δεν ικανοποιείται η συνθήκη τερματισμού
   1. Για κάθε πρότυπο
      1. Υπολογισμός της εξόδου του δικτύου
      2. Εκπαίδευση των βαρών.
   2. Στο τέλος κάθε εποχής, υπολογισμός του σφάλματος εκπαίδευσης

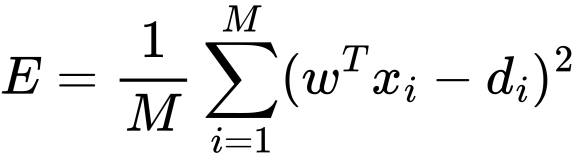
ΤΕΡΜΑΤΙΣΜΟΣ:  
 Αν δεν έχει χρειαστεί εκπαίδευση των βαρών για μία ολόκληρη εποχή ή

αν έχουμε ξεπεράσει τον μέγιστον αριθμό επαναλήψεων - εφόσον αυτός έχει οριστεί.

ΕΞΟΔΟΣ:

Τα ενημερωμένα βάρη

Αρχικά, πρέπει να αποθηκευτούν τα διανύσματα προτύπων που έχουμε στη διάθεση μας μαζί με τις αντίστοιχες επιθυμητές εξόδους. Και σε αυτή την περίπτωση - όπως και στον Perceptron, μπορεί να προκύψουν μη αξιόπιστα αποτελέσματα αν υπάρχει κάποιο λάθος στα πρότυπα ή στις επιθυμητές εξόδους τους. Το δεύτερο βήμα είναι η αρχικοποίηση του ρυθμού μάθησης καθώς και του διανύσματος βαρών που θα χρησιμοποιηθεί έτσι ώστε να εκπαιδευτεί κατάλληλα το δίκτυο, προκειμένου να έχουμε το επιθυμητό αποτέλεσμα. Η αρχικοποίηση του ρυθμού μάθησης είναι ίδια με αυτή του Perceptron, δηλαδή σε μία μικρή θετική δεκαδική τιμή, συνήθως στο κλειστό διάστημα [0,1]. Συχνά προτιμούμε την τιμή 0.1. (Τσούλος, χ.χ.) Οι τρόποι αρχικοποίησης των τιμών των βαρών, είναι ίδιοι με αυτούς κατά την υλοποίηση ενός Perceptron δικτύου. Θα πρέπει πρώτα να γνωρίζουμε το πλήθος των τιμών που χρειαζόμαστε, το οποίο ισούται με το πλήθος των χαρακτηριστικών που έχουμε στη διάθεση μας αυξημένο κατά 1 για το bias. Ο πιο απλός τρόπος αρχικοποίησης είναι να θέσουμε όλες τις τιμές είτε σε μία τιμή ίση με το μηδέν (0) είτε με την μονάδα (1). Εναλλακτικά, μπορεί να γίνει αρχικοποίηση των τιμών σε τυχαίες τιμές στο κλειστό διάστημα [0,1]. Στη συνέχεια γίνεται μία επαναληπτική διαδικασία για κάθε πρότυπο που έχουμε στην διάθεση μας. Έτσι, για κάθε πρότυπο υπολογίζουμε την έξοδο του δικτύου χρησιμοποιώντας τις τιμές των βαρών που έχουμε στη διάθεση μας.. Από εδώ και πέρα ξεκινάει η διαφοροποίηση του Adaline με το Perceptron. Ο υπολογισμός αυτός - σε αντίθεση με το Perceptron, δε προκύπτει από την συνάρτηση ενεργοποίησης, αλλά από το γινόμενο ανάμεσα στην είσοδο και τα συναπτικά βάρη. Σε κάθε περίπτωση γίνεται εκπαίδευση των βαρών και με αυτόν τον τρόπο, ενσωματώνεται και η έννοια του χρόνου. Στο τέλος κάθε εποχής, θα πρέπει να υπολογιστεί το σφάλμα εκπαίδευσης και συνήθως υπολογίζεται βάσει του μέσου τετραγωνικού σφάλματος που προκύπτει από τον παρακάτω τύπο:

[[6]](#footnote-5)

Η διαδικασία αυτή τελειώνει είτε όταν έχουμε μία ολόκληρη εποχή κατά την οποία παρά την εκπαίδευση των βαρών, δεν γίνεται κάποια μεταβολή των τιμών τους είτε μετά από κάποιο πεπερασμένο αριθμό επαναλήψεων - εφόσον έχει οριστεί στην αρχή της διαδικασίας. Η ενημέρωση των βαρών γίνεται με την χρήση του ίδιου τύπου που χρησιμοποιούμε καισ το δίκτυο Perceptron.

## 2.3 Τα δίκτυα MLP

Όπως αναφέρθηκε σε προηγούμενα υποκεφάλαια, το δίκτυο Perceptron μπορεί να χρησιμοποιηθεί για τον διαχωρισμό προτύπων μόνο δύο κατηγοριών εφόσον αυτά είναι γραμμικά διαχωρίσιμα, ενώ το δίκτυo Adaline μπορεί να είναι αποδοτικό σε πρότυπα δύο κατηγοριών ακόμη και αν αυτά είναι σχεδόν γραμμικά διαχωρίσιμα. Εξέλιξη τους λοιπόν, είναι τα δίκτυα MLP, τα οποία αποτελούνται από περισσότερους από έναν νευρώνες, αίροντας τους περιορισμούς που έχουν αποδειχθεί για το Perceptron. 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 στα κρυφά στρώματα, καθώς αποδεικνύεται πως με την χρήση της, τα δίκτυα αυτά μπορούν να προσεγγίσουν οποιαδήποτε ομαλή συνάρτηση, όσο κοντά επιθυμούμε. Αυτός είναι και ο λόγος που αυτά τα δίκτυα καλούνται και Universal Approximators (Καθολικοί Προσεγγιστές). Το στρώμα εξόδου συνηθίζεται να χρησιμοποιεί τη βηματική συνάρτηση. 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, αποκλειστική διάζευξη). (Haykin, 2010)· (Διαμαντάρας, 2007)· (Τσούλος, χ.χ)

Μπορούμε να πούμε πως η θεωρητική ανάλυση αυτών των δικτύων είναι αρκετά δύσκολη υπόθεση. Αυτό οφείλεται λόγω των κρυφών στρωμάτων 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 διαδίδονται 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 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 διαφοροποιεί τα MLP δίκτυα από το δίκτυο Perceptron του Rosenblatt. (Haykin, 2010)

Μπορούμε να παρατηρήσουμε στο παρακάτω σχήμα ένα δίκτυο MLP με τρεις (3) νευρώνες εισόδου στο στρώμα εισόδου, ένα (1) κρυμμένο στρώμα με τέσσερις (4) νευρώνες επεξεργασίας και στο στρώμα εξόδου τρεις (3) εξόδους. Συνηθίζεται να έχουμε στο στρώμα εξόδου τόσους νευρώνες όσες και οι διαφορετικές κατηγορίες που έχουμε στη διάθεση μας.



(Εικόνα 2.2) MLP

Η υλοποίηση ενός δικτύου MLP, περιγράφεται με την παρακάτω μορφή ψευδοκώδικα:

ΕΙΣΟΔΟΙ:   
 Τα πρότυπα εισόδου με τις αντίστοιχες επιθυμητές εξόδους  
  
ΜΕΘΟΔΟΣ:

1. Αρχικοποίηση των βαρών που θα χρησιμοποιηθούν
2. Αρχικοποίηση του ρυθμού μάθησης
3. Αρχικοποίηση του πλήθους των νευρώνων που θα χρησιμοποιηθούν
4. Αρχικοποίηση κατωφλιού προκειμένου να τερματίσει ο αλγόριθμος
5. Αρχικοποίηση του μέγιστου αριθμού επαναλήψεων - προαιρετικό
6. Όσο δεν ικανοποιείται η συνθήκη τερματισμού
   1. Για κάθε πρότυπο
      1. Υπολογισμός της εξόδου του δικτύου
      2. Υπολογισμός του σφάλματος του δικτύου
      3. Εκπαίδευση των βαρών

ΤΕΡΜΑΤΙΣΜΟΣ:  
Εφόσον το σφάλμα εκπαίδευσης μειωθεί κάτω από το προκαθορισμένο κατώφλι ή ξεπεραστεί ο μέγιστος αριθμός επαναλήψεων - εφόσον έχει οριστεί.

ΕΞΟΔΟΣ:

Τα ενημερωμένα βάρη

Αρχικά, πρέπει να αποθηκευτούν τα διανύσματα προτύπων που έχουμε στη διάθεση μας μαζί με τις αντίστοιχες επιθυμητές εξόδους. Και σε αυτή την περίπτωση - όπως στα δίκτυα Perceptron και Adaline, μπορεί να προκύψουν μη αξιόπιστα αποτελέσματα αν υπάρχει κάποιο λάθος στα πρότυπα ή στις επιθυμητές εξόδους τους. Το δεύτερο βήμα είναι η αρχικοποίηση της τιμής του ρυθμού μάθησης που έχουμε επιλέξει για αυτό το δίκτυο καθώς και του πλήθους των κόμβων που έχουμε καταλήξει πως επιθυμούμε να χρησιμοποιήσουμε. Επιπλέον, πρέπει να γίνει αρχικοποίηση του διανύσματος βαρών που θα χρησιμοποιηθεί έτσι ώστε το δίκτυο να εκπαιδευτεί κατάλληλα προκειμένου να έχουμε το επιθυμητό αποτέλεσμα. Οι τρόποι αρχικοποίησης των τιμών των βαρών, είναι ίδιοι με αυτούς που χρησιμοποιούνται στα δίκτυα Perceptron και Adaline. Ο πιο απλός είναι αυτός της αρχικοποίησης όλων των τιμών σε μία τιμή είτε ίση με το μηδέν (0) είτε με την μονάδα (1). Εναλλακτικά, μπορεί να γίνει αρχικοποίηση σε τυχαίες τιμές στο κλειστό διάστημα [0,1]. Το σύνολο των τιμών που θέλουμε να έχουμε για τα βάρη προκύπτει από τον τύπο: (d + 2) κατά τον οποίο, d είναι η διάσταση των χαρακτηριστικών των προτύπων και H το πλήθος των κόμβων που έχουμε επιλέξει να χρησιμοποιήσουμε για το δίκτυο. Μία ακόμη σημαντική παράμετρος που πρέπει να αρχικοποιηθεί είναι η τιμή του κατωφλιού, ώστε σε περίπτωση που το σφάλμα εκπαίδευσης του δικτύου έχει χαμηλότερη τιμή από αυτό, να τερματίζει ο αλγόριθμος. Στη συνέχεια γίνεται μία επαναληπτική διαδικασία για κάθε πρότυπο που έχουμε στην διάθεση μας. Έτσι, για κάθε πρότυπο υπολογίζουμε την έξοδο του δικτύου χρησιμοποιώντας τις τιμές των βαρών που έχουμε στη διάθεση μας. Σε αυτή την περίπτωση δικτύου υπολογίζουμε την έξοδο με τη χρήση κατάλληλων συναρτήσεων ενεργοποίησης μέσω της προώθησης του προτύπου από το ένα στρώμα του δικτύου στο αμέσως επόμενο. Εφόσον ξέρουμε την έξοδο του δικτύου για το πρότυπο, επόμενο βήμα είναι ο υπολογισμός του σφάλματος του δικτύου που προκύπτει από την διαφορά ανάμεσα στην επιθυμητή και την πραγματική τιμή της εξόδου. Σε κάθε περίπτωση γίνεται εκπαίδευση των βαρών, χρησιμοποιώντας τον αλγόριθμο εκπαίδευσης Back Propagation. Θα γίνει ανάλυση του, στο αμέσως επόμενο υποκεφάλαιο. Η διαδικασία αυτή τελειώνει είτε όταν το σφάλμα εκπαίδευσης έχει τιμή χαμηλότερη από το κατώφλι που έχουμε αρχικοποιήσει στην αρχή της υλοποίησης είτε μετά από κάποιο πεπερασμένο αριθμό επαναλήψεων - εφόσον έχει οριστεί στην αρχή της διαδικασίας.

## 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 κατηγοριοποιήσουν μόνο γραμμικά διαχωρίσιμα πρότυπα. Κυριότερος εκπρόσωπος των αλγορίθμων εκπαίδευσης MLP είναι ο Back Propagation (Οπισθοδιάδοσης σφάλματος). Βασίζεται στο γενικευμένο κανόνα Δέλτα (generalized Delta rule) ο οποίος επιτρέπει να καθοριστεί το ποσοστό του συνολικού σφάλματος που αντιστοιχεί στα βάρη του κάθε νευρώνα, ακόμη και αυτών που ανήκουν στα κρυφά επίπεδα, για τους οποίους η επιθυμητή έξοδος δεν είναι γνωστή. Ο απλός κανόνας Δέλτα δεν είναι σε θέση να λειτουργήσει σε κρυφά στρώματα. Έτσι, είναι δυνατό να υπολογιστούν οι διορθώσεις στα βάρη του κάθε νευρώνα ξεχωριστά. O Back Propagation περιλαμβάνει τον LMS σαν ειδική περίπτωση. Η χρήση του όρου Back Propagation εμφανίζεται στο προσκήνιο μετά από το 1985, όταν ο όρος απέκτησε ευρεία απήχηση χάρη στο σημαντικό βιβλίο με τίτλο Parallel Distributed Processing.

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. Ιδανικό σενάριο θεωρούμε την περίπτωση που υπάρχει πλήρη ταύτιση ανάμεσα στις πραγματικές και επιθυμητές εξόδους. Ωστόσο αυτό μπορεί να είναι απολύτως εφικτό, και για αυτόν τον λόγο επιζητούμε τη βέλτιστη προσέγγιση της επιθυμητής κατάστασης χρησιμοποιώντας ένα κριτήριο κόστους. Κλασικό παράδειγμα κριτηρίου κόστους αποτελεί το μέσο τετραγωνικό σφάλμα., το οποίο είδαμε να χρησιμοποιείται και στα δίκτυα Adaline. (Βλαχάβας et al., 2006)· (Διαμαντάρας, 2007) 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 όπως η αργή σύγκλιση αλλά και το ότι δεν υπάρχει και εγγύηση γενίκευσης. (Δούνιας και Καραμπότσης, χ.χ.)

Παρακάτω υπάρχει με την μορφή ψευδοκώδικα μία πιο αναλυτική μορφή του δικτύου MLP με τη χρήση της μεθόδου Back Propagation προκειμένου να γίνει η εκπαίδευση βαρών. Ουσιαστικά, στο σημείο που πρέπει να γίνει η εκπαίδευση βαρών σε ένα δίκτυο MLP μπορούμε να καταλάβουμε ότι χρειάζονται δύο επιπλέον βήματα. Το πρώτο είναι ο υπολογισμός των τοπικών παραγώγων σφάλματος για τα στρώματα εξόδου. Το δεύτερο είναι ο υπολογισμός των τοπικών παραγώγων σφάλματος για τα κρυφά στρώματα, ξεκινώντας από το πίσω προς το πιο μπροστά στρώμα. Αφού γίνουν αυτά τα δύο βήματα, προκειμένου να γίνει η εκπαίδευση των βαρών, θα πρέπει να γίνει χρήση μίας ακόμη μεθόδου που ονομάζεται Gradient Descent. Θα παρουσιαστεί αναλυτικά στο αμέσως επόμενο υποκεφάλαιο.

ΕΙΣΟΔΟΙ:   
 Τα πρότυπα εισόδου με τις αντίστοιχες επιθυμητές εξόδους  
  
ΜΕΘΟΔΟΣ:

1. Αρχικοποίηση των βαρών που θα χρησιμοποιηθούν
2. Αρχικοποίηση του ρυθμού μάθησης
3. Αρχικοποίηση του πλήθους των νευρώνων που θα χρησιμοποιηθούν
4. Αρχικοποίηση κατωφλιού προκειμένου να τερματίσει ο αλγόριθμος
5. Αρχικοποίηση του μέγιστου αριθμού επαναλήψεων - προαιρετικό
6. Για κάθε πρότυπο
   1. Υπολογισμός της εξόδου του δικτύου
   2. Υπολογισμός του σφάλματος του δικτύου
   3. Χρήση της μεθόδου Back Propagation
      1. Υπολογισμός των τοπικών παραγώγων σφάλματος για τα στρώματα εξόδου
      2. Υπολογισμός των τοπικών παραγώγων σφάλματος για τα κρυφά στρώματα
      3. Ενημέρωση των βαρών

ΤΕΡΜΑΤΙΣΜΟΣ:  
 Εφόσον το σφάλμα εκπαίδευσης μειωθεί κάτω από κάποιο προκαθορισμένο κατώφλι ή ξεπεραστεί ο μέγιστος αριθμός επαναλήψεων - εφόσον έχει οριστεί.

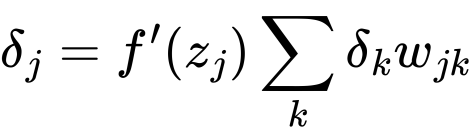
ΕΞΟΔΟΣ:

Τα ενημερωμένα βάρη

Ο υπολογισμός των τοπικών παραγώγων σφάλματος διαφέρει για τα στρώματα εξόδου και τα κρυφά στρώματα. Ο λόγος είναι πως στα κρυφά στρώματα δε γνωρίζουμε την έξοδο. Έτσι προκύπτει ο παρακάτω μαθηματικός τύπος για τον υπολογισμό των τοπικών παραγώγων σφάλματος για τα στρώματα εξόδου

C:/Users/Evita/AppData/Local/Temp/wps.vgmvwnwps[[7]](#footnote-6)

Και ο παρακάτω μαθηματικός τύπος για τον υπολογισμό των τοπικών παραγώγων σφάλματος για τα κρυφά στρώματα.

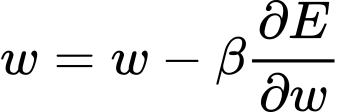
[[8]](#footnote-7)

## 2.5 Η μέθοδος Gradient Descent

Η μέθοδος Gradient Descent ή αλλιώς μέθοδος κατάβασης δυναμικού, παίρνει και το όνομα “Γενικευμένος Κανόνας Δέλτα” σύμφωνα με τον Διαμαντάρα. Είναι μία από τις πιο δημοφιλείς μεθόδους για την εκτέλεση βελτιστοποίησης (Sebastian, χ.χ.) και είναι μακράν ο πιο συνηθισμένος τρόπος για τη βελτιστοποίηση νευρωνικών δικτύων. Είναι μια κλασική μέθοδος εύρεσης της ελάχιστης τιμής μίας συνάρτησης κόστους κάποιου πλήθoυς 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). (Διαμαντάρας, 2007) Μπορεί επίσης να αποδοθεί ως επικλινής κάθοδος ή επικλινής κατάδυση. (Θεοδωρίδης και Κουτρούμπας, 2012)·

Υπάρχουν τρεις παραλλαγές gradient descent, οι οποίες διαφέρουν ανάλογα με τον όγκο των δεδομένων που χρησιμοποιούμε για τον υπολογισμό της κλίσης της αντικειμενικής συνάρτησης κάνοντας επιλογή ανάμεσα στην ακρίβεια ενημέρωσης των παραμέτρων και του χρόνο που θα πραγματοποιηθεί μία ενημέρωση. Η πρώτη επιλογή είναι αυτή της μεθόδου Batch gradient descent κατά την οποία υπολογίζει την κατάβαση δυναμικού στο τέλος μίας εποχής. (Sebastian, χ.χ.) Η δεύτερη επιλογή είναι αυτή της μεθόδου Stochastic Gradient Descent κατά την οποία ο υπολογισμός γίνεται κάθε φορά που έρχεται ένα πρότυπο και η τρίτη επιλογή χρησιμοποιεί έναν συνδυασμό των παραπάνω δύο, προκειμένου η εποχή να μοιραστεί σε μικρότερα σύνολα προτύπων, στο τέλος των οποίων θα γίνουν οι απαραίτητοι υπολογισμοί.

Κατά την εκπαίδευση βαρών σε ένα δίκτυο MLP με τη χρήση των μεθόδων Back Propagation και Gradient Descent (βλέπε ψευδοκώδικα Back Propagation) γίνεται χρήση του παρακάτω μαθηματικού τύπου:

[[9]](#footnote-8)

## 2.6 Εφαρμογές νευρωνικών δικτύων

Τα ΤΝΔ βρίσκουν πληθώρα εφαρμογών σε διαφορετικούς τομείς της καθημερινότητας μας. 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 είτε προερχόμενων από συσκευές radar με σκοπό την ανίχνευση και την παρακολούθηση αντικειμένων στον αέρα, στη γη, ακόμη και στο νερό είτε προερχόμενων από συσκευές sonar με σκοπό την ανίχνευση και τον εντοπισμό αντικειμένων κάτω από το νερό. 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.

Μία ακόμη κατηγορία προβλημάτων είναι τα προβλήματα αποτίμησης (assessment). 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. Στη μηχανολογία χρησιμοποιούνται εφαρμογές παρακολούθησης, επιθεώρησης και ελέγχου των προϊόντων.

Τέλος, στον τομέα της πρόβλεψης, ενσωματώνονται εφαρμογές πρόβλεψης ισοτιμίας νομισμάτων, τιμών μετοχών, πωλήσεων, αλλά 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 στον χώρο καθώς απαιτεί πρόβλεψη της κίνησης στον χώρο μέσω δυναμικών εμποδίων. Μπορούμε να αναφέρουμε επιπλέον τη λήψη αποφάσεων, διότι περιλαμβάνει την πρόβλεψη των συνεπειών διάφορων επιλογών. Τέλος, αξίζει να αναφερθεί η πρόβλεψη οικονομικών μεγεθών όπως η πτώχευση και οι 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 αντιμετώπισή του με ΤΝΔ απαιτεί μελέτη και πειραματισμό. (Τσούλος, χ.χ.) & (Βλαχάβας et al., 2006) & (Ντούνης, 2024).

# Γενετικοί αλγόριθμοι

## Ιστορική Αναδρομή

Σημαντικό ρόλο για την ανάπτυξη των γενετικών αλγορίθμων αποτέλεσε ο Friedberg το 1958, όπου επιχείρησε την αυτόματη παραγωγή σύνθετων προγραμμάτων FORTRAN με το συνδυασμό μικρότερων προγραμμάτων - αν και αυτά τις περισσότερες φορές δεν ήταν εκτελέσιμα. Οι γενετικοί αλγόριθμοι αναπτύχθηκαν από τον John Holland τους συναδέλφους του αλλά και τους μαθητές του, στο πανεπιστήμιο του Michigan. Αποτελούν τους πιο δημοφιλής εξελικτικούς αλγόριθμους και βασίζονται στη θεωρία της εξέλιξης 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 σύμφωνα με τους (Βλαχάβα et al) αποτελούν μία ενδιαφέρουσα κατηγορία πιθανοκρατικών αλγορίθμων. Αυτοί οι αλγόριθμοι χρησιμοποιούνται για την επίλυση προβλημάτων που το μέγεθος τους καθιστά απαγορευτική τη χρήση κλασσικών μεθόδων αναζήτησης. Τα προβλήματα αυτά είναι ειδικότερα προβλήματα βελτιστοποίησης (optimization) και υπάρχουν κυρίως σε βιομηχανικές εφαρμογές όπως στη σχεδίαση VLSI κυκλωμάτων. (Βλαχάβας et al., 2006) & (Καμπουρλάζος & Παπακώστας, 2015)

## 3.2 Μέθοδοι κωδικοποίησης

Στην κλασσική προσέγγιση των γενετικών αλγορίθμων, κάθε χρωμόσωμα αναπαρίσταται με μία συμβολοσειρά ενός πεπερασμένου αλφαβήτου. Συνήθως χρησιμοποιείται το δυαδικό αλφάβητο, για αυτό οι συμβολοσειρές ονομάζονται και δυαδικές συμβολοσειρές (bit-strings). Ωστόσο υπάρχουν περιπτώσεις που χρησιμοποιούν και πιο σύνθετες μορφές αναπαράστασης.

Προκειμένου να γίνει αρχικοποίηση του πληθυσμού χρειάζεται να παρθεί απόφαση για την κωδικοποίηση που θα χρησιμοποιηθεί. Αυτή εξαρτάται από τη φύση του προβλήματος καθώς και από τη μορφή των παραμέτρων του. Μερικές μέθοδοι είναι η δυαδική κωδικοποίηση (binary encoding), η κωδικοποίηση μετάθεσης (permutation encoding), η κωδικοποίηση δέντρου (tree encoding) καθώς και η κωδικοποίηση τιμών (value encoding). (Καμπουρλάζος & Παπακώστας, 2015)

## 3.3 Συνάρτηση καταλληλότητας

Η συνάρτηση καταλληλότητας αποτελεί το κριτήριο για την αξιολόγηση των χρωμοσωμάτων. Η συνάρτηση δέχεται ως είσοδο ένα χρωμόσωμα και επιστρέφει έναν αριθμό που υποδηλώνει το βαθμό καταλληλότητας του. Το πεδίο τιμών της συνάρτησης είναι συνήθως το σύνολο των πραγματικών αριθμών στο κλειστό διάστημα [0,1]. Σε αυτή την περίπτωση η τιμή 1 υποδηλώνει ότι το χρωμόσωμα ικανοποιεί όλες τις προϋποθέσεις του προβλήματος και αποτελεί αποδεκτή λύση ενώ η τιμή 0 υποδηλώνει μη αποδεκτή λύση.

Ο γενικός κανόνας για την κατασκευή της συνάρτησης είναι πως πρέπει να αντικατοπτρίζει την αξία του χρωμοσώματος. Ωστόσο υπάρχουν περιπτώσεις στις οποίες δεν έχει τόσο νόημα η αξία του χρωμοσώματος, αλλά χρειαζόμαστε μία τιμή για το πόσο κοντά βρισκόμαστε σε μία επιθυμητή λύση.

Μία προσέγγιση που ακολουθείται πολλές φορές είναι αυτή την προσεγγιστικής συνάρτησης καταλληλότητας (approximate fitness function). Το κρίσιμο θέμα είναι η επιθυμητή ακρίβεια της συνάρτησης καταλληλότητας και το υπολογιστικό κόστος που θεωρείται αποδεκτό για μία συνάρτηση που ενδεχομένως δίνει άριστα αποτελέσματα. Για παράδειγμα, σε ένα πρόβλημα μπορεί να υπάρχουν διαθέσιμες δύο συναρτήσεις καταλληλότητας, μία ακριβής και μία προσεγγιστική. Αν η ακριβής συνάρτηση έχει δεκαπλάσιο χρόνο υπολογισμού από την προσεγγιστική, ενδεχομένως να είναι αποδοτικότερη η προσεγγιστική συνάρτηση καθώς η διαδικασία αξιολόγησης εκτελείται πάρα πολλές φορές και έτσι επιτρέπει την πραγματοποίηση δεκαπλάσιας ποσότητας υπολογισμών στον ίδιο χρόνο σε σχέση με την ακριβής συνάρτηση καταλληλότητας. (Βλαχάβας et al., 2006)

## 3.4 Γενετικοί τελεστές

Οι ζωντανοί οργανισμοί εξελίσσονται χρησιμοποιώντας δύο βασικούς τελεστές ή αλλιώς μηχανισμούς. Ο πρώτος αναφέρεται ως φυσική επιλογή (natural selection) και καθορίζει ποια άτομα του πληθυσμού θα επιβιώσουν με σκοπό την αναπαραγωγή. Ο δεύτερος αναφέρεται ως διασταύρωση (crossover) και σκοπός του είναι η εξασφάλιση της ανάμειξης των γονιδίων (genes) των γονέων ώστε να δημιουργηθεί μία νέα γενιά χρωμοσωμάτων με το συνδυασμό των χαρακτηριστικών που έχουν οι γονείς.

Η διαδικασία επιλογής των χρωμοσωμάτων - γονέων σχετίζεται με την απόδοση πιθανοτήτων επιλογής προς αναπαραγωγή στα μέλη ενός πληθυσμού. Κατά τη διαδικασία αυτή, κάποια χρωμοσώματα με υψηλή τιμή στη συνάρτηση καταλληλότητας ενδέχεται να επιλεγούν προς αναπαραγωγή περισσότερες από μία φορές, ενώ κάποια άλλα με χαμηλή καταλληλότητα ενδέχεται να μην επιλεγούν καθόλου. ‘Έχουν προταθεί και αναπτυχθεί πολλές μέθοδοι επιλογής οι οποίες κατηγοριοποιούνται σε στοχαστικές (stochastic) και αιτιοκρατικές (deterministic). Στις αιτιοκρατικές μεθόδους αξίζει να αναφερθούμε στην επιλογή του καλύτερου συζύγου (top mate selection) που αποτελεί και την πιο απλή μέθοδο καθώς και σε αυτή της τοπικής επιλογής (local selection). Υπολογιστικός πειραματισμός έδειξε πως η επιλογή των γονέων δεν πρέπει να βασίζεται αποκλειστικά στα καλύτερα χρωμοσώματα διότι κάτι τέτοιο μπορεί να αποκλείσει τη βέλτιστη λύση. Μία καλή πρακτική είναι να επιζούν λιγότερα καλά χρωμοσώματα, ώστε να διασφαλίζεται η μετάδοση της γενετικής τους πληροφορίας. Στις στοχαστικές μεθόδους, ανήκει η πιθανολογική καθολική δειγματοληψία (stohastic universal sampling), η επιλογή ρουλέτας (roulette wheel selection) καθώς και η επιλογή με διαγωνισμό (tournament selection).

Σε αυτή την πτυχιακή εργασία έχει επιλεχθεί ο πιο συνηθισμένος τελεστής επιλογής, που είναι η επιλογή της ρουλέτας. Σε αυτή τη μέθοδο όσο καλύτερα είναι τα χρωμοσώματα, τόσο περισσότερες πιθανότητες έχουν να επιλεγούν. Βέβαια, ενδέχεται να επιλεγούν για αναπαραγωγή χρωμοσώματα με κακή τιμή της συνάρτησης καταλληλότητας. Η λογική του παραπάνω αλγορίθμου είναι ότι οι υποψήφιες λύσεις με μεγάλη τιμή καταλληλότητας έχουν μεγαλύτερη πιθανότητα να αυξήσουν την τιμή του καταχωρητή Κ ώστε να υπερβεί την τιμή n και συνεπώς να επιλεγούν. Επίσης, κάποιες υποψήφιες λύσεις μπορεί να επιλεγούν περισσότερο από μία φορές. Μετά την ολοκλήρωση της παραπάνω διαδικασίας, από τα μέλη της δεξαμενής ζευγαρώματος δημιουργούνται ζευγάρια με τυχαίο τρόπο και οι απόγονοι που προκύπτουν αποτελούν το νέο πληθυσμό.

Από την στιγμή που έχουν επιλεχτεί τα χρωμοσώματα γονείς, γίνεται η αναπαραγωγή, δηλαδή η διαδικασία δημιουργίας απογόνων. Η αναπαραγωγή γίνεται με τη βοήθεια του τελεστή της διασταύρωσης. Έτσι, από δύο γονείς, αντιγράφοντας επιλεγμένα bit από κάθε γονέα με τρόπο τέτοιο ώστε το i-οστό bit του απογόνου να αντιστοιχεί στο i-οστό bit ενός εκ των γονέων του. Το ποιος γονέας θα συνεισφέρει το κάθε bit αποφασίζεται βάσει ενός μηχανισμού που ονομάζεται μάσκα διασταύρωσης (crossover mask). Η διασταύρωση δεν εφαρμόζεται σε όλα τα ζεύγη των επιλεγμένων για αναπαραγωγή ατόμων, αλλά γίνεται μία τυχαία επιλογή. Αν η διασταύρωση δεν εφαρμοστεί, τότε οι γονείς αντιγράφονται στην επόμενη γενεά. Υπάρχουν διάφορες μέθοδοι διασταύρωσης. Κριτήριο για την επιλογή της μεθόδου που θα χρησιμοποιηθεί αποτελεί τόσο το είδος του προβλήματος, όσο και το είδος της εφαρμοζόμενης κωδικοποίησης. Αξίζει να αναφερθεί η διασταύρωση ενός σημείου (single-point crossover), η διασταύρωση δύο σημείων (two-point crossover), διασταύρωση πολλαπλών σημείων (multi-point crossover), η αριθμητική διασταύρωση (arithmetic crossover), καθώς και η ομοιόμορφη διασταύρωση (uniform crossover).

Σε αυτή την πτυχιακή εργασία έχει επιλεχθεί η υλοποίηση της διασταύρωσης δύο σημείων. Σκοπός αυτής της μεθόδου είναι ο γονέας να τεμαχίζεται σε Ν σημεία, στην δικιά μας περίπτωση σε δύο (2). Έπειτα, ανταλλάσσονται αμοιβαία τμήματα των χρωμοσωμάτων των γονέων, τα οποία προέκυψαν από τη διαίρεση των χρωμοσωμάτων τους στα σημεία διασταύρωσης. Η ιδέα πίσω από αυτή την μέθοδο είναι πως τα τμήματα ενός χρωμοσώματος που συμβάλλουν περισσότερο στην απόδοση του ενδέχεται να μην είναι γειτονικά. Έτσι, ανταλλάσσονται τμήματα των χρωμοσωμάτων με την προσδοκία βελτίωσης της απόδοσης ενός χρωμοσώματος. Υπολογιστικά πειράματα συχνά επαληθεύουν την προσδοκία αυτήν.

Ένας ακόμη τελεστής των γενετικών αλγορίθμων είναι αυτός της μετάλλαξης (mutation). Πρόκειται για μία διαδικασία που εμφανίζεται στους ζωντανούς οργανισμούς βάσει της οποίας παρατηρείται μία τυχαία αλλαγή σε κάποια χρωμοσώματα υπό ειδικές συνθήκες. Στους γενετικούς τελεστές η εφαρμογή της μετάλλαξης γίνεται με τυχαιότητα τόσο για το χρωμόσωμα όσο και γ ια το γονίδιο στο οποίο θα εφαρμοστεί. Σκοπός αυτού του τελεστή είναι η διασφάλιση της ποικιλομορφίας του πληθυσμού και αυτό επιτυγχάνεται με το να επιτρέπει τον τυχαίο σχηματισμό χρωμοσωμάτων. Έτσι, συμβάλλει στη διεύρυνση της εξερεύνησης του χώρου των λύσεων.

Τέλος, ένας ακόμη τελεστής είναι ο ελιτισμός (elitism), κατά τον οποίο βάσει ενός προκαθορισμένου αριθμού χρωμοσωμάτων που κρίνονται ως “πλέον κατάλληλα”, επαναλαμβάνονται αυτούσια στο νέο πληθυσμό. Έτσι, εξασφαλίζεται πως κάποια “πλέον κατάλληλα” χρωμοσώματα θα δε εξαφανιστούν ως αποτέλεσμα της εφαρμογής των τελεστών που αναφέρθηκαν παραπάνω. (Καμπουρλάζος & Παπακώστας, 2015) & (Βλαχάβας et al., 2006)

## 3.5 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.

Έχουν εφαρμοστεί επιτυχώς σε προβλήματα βελτιστοποίησης όπως δρομολόγηση καλωδίων (wire routing), χρονοπρογραμματισμό (scheduling), προσαρμοστκό έλεγχο (adaptive control), παίγνια (game playing), γνωστική μοντελοποίηση (cognitive modeling), προβλήματα εφοδιαστικής (logistic), προβλήματα πλανώδιου πωλητή, προβλήματα βέλτιστου ελέγχου, βελτιστοποίηση ερωτήσεων σε Βάσεις Δεδομένων και άλλα.

Το σύστημα GABIL, είναι ένα χαρακτηριστικό παράδειγμα χρήσης των γενετικών αλγορίθμων με σκοπό τη μάθηση Boolean εννοιών που αναπαρίστανται από ένα διαζευκτικό (disjunctive) σύνολο προτασιακών κανόνων. Στο σύστημα αυτό, χρησιμοποιήθηκαν τεχνητά προβλήματα μάθησης προκειμένου να εξεταστεί η ακρίβεια του, εφαρμόστηκε όμως και σε πραγματικό πρόβλημα, τη διάγνωση καρκίνου του μαστού. (Βλαχάβας et al., 2006)

# Μέθοδος - αποτελέσματα

## 4.1 Τα dataset που χρησιμοποιήθηκαν

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

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

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

Πίνακας 3.1 [Τίτλος πίνακα]

## 4.2 Η προτεινόμενη μέθοδος

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

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

Πίνακας 3.2 [Τίτλος πίνακα]

## 4.3 Πειραματικά αποτελέσματα

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

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

Πίνακας 3.2 [Τίτλος πίνακα]

# Συμπεράσματα

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

# ΒΙΒΛΙΟΓΡΑΦΙΑ

* Anderson, D., & 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.* Ανακτήθηκε στις 08/06/2024 από <https://www.imada.sdu.dk/u/rolf/Edu/DM534/E16/DM534-marco.pdf>
* ~~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.
* Sebastian Ruder (χ.χ.) *An overview of gradient descent optimization algorithms. Ανακτήθηκε στις 22/08/2024 από* <https://arxiv.org/pdf/1609.04747>
* Suzuki, K., (ed.) (2011) *Artificial Neural Networks - Methodological advances and biomedical applications*. Croatia: Published by InTech
* Βλαχάβας, Ι., Κεφάλας, Π., Βασιλειάδης, Ν., Κόκκορας, Φ. & Σακελλαρίου, Η., (2006) *Τεχνητή Νοημοσύνη*. 3η έκδοση. Θεσσαλονίκη: Εκδόσεις Πανεπιστημίου Μακεδονίας.
* Γεωργούλη, Κ., (2015) *Τεχνητή Νοημοσύνη - Μία εισαγωγική προσέγγιση*. Αθήνα: Κάλλιπος Ανακτήθηκε στις 13/07/2024 από <https://repository.kallipos.gr/handle/11419/3381>
* Διαμαντάρας, Κ., (2007) *Τεχνητά Νευρωνικά Δίκτυα*. Αθήνα: Εκδόσεις Κλειδάριθμος.
* Δούνιας, Γ., και Καραμπότσης, Ε., (χ.χ.) *Εισαγωγή στα Τεχνητά Νευρωνικά Δίκτυα.* Ανακτήθηκε στις 15/06/2024 από <https://mde-lab.aegean.gr/images/stories/docs/ANNs-intro.pdf>
* Ηλιάδης, Λ., (χ.χ) *Νευρωνικά Δίκτυα*. Ανακτήθηκε στις 01/07/2024 από <https://slideplayer.gr/slide/7018453/>
* Θεοδωρίδης, Σ., & Κουτρούμπας, Κ. ,(2012) *Αναγνώριση προτύπων*. Κύπρος: Εκδόσεις Πασχαλίδη.
* Καμπουρλάζος, Β. Γ. & Παπακώστας, Γ. Α., (2015) *Εισαγωγή στην υπολογιστική νοημοσύνη*. Αθήνα: Κάλλιππος.
* Κίτρου, Κ. (χ.χ.) *Τι είναι μηχανική μάθηση; (Machine Learning).* Ανακτήθηκε στις 08/06/2024 από
* Λύκας, Α., (χ.χ.) *Βασικές αρχές εκπαίδευσης ΤΝΔ: Το perceptron*. Ανακτήθηκε στις 10/06/2024 από <https://www.cs.uoi.gr/~arly/courses/nn/slides/K2.pdf>
* Λυκοθανάσης, Σ. & Κουτσομητρόπουλος, Δ., (2021) *Υπολογιστική νοημοσύνη και βαθιά μάθηση*. Εκδόσεις Κάλλιπος Ανακτήθηκε στις 13/07/2024 από <https://repository.kallipos.gr/handle/11419/9117>
* Ντούνης, Α., (2024) *Βασικές Αρχές Υπολογιστικής Νοημοσύνης*. Εκδόσεις Κάλλιπος Ανακτήθηκε στις 12/07/2024 από <https://repository.kallipos.gr/handle/11419/13341>
* Πούλος, Μ., (2015) *Σημασιολογική Επεξεργασία της Πληροφορίας*. Εκδόσεις Κάλλιπος Ανακτήθηκε στις 12/07/2024 από <https://repository.kallipos.gr/handle/11419/2854>
* Τσίπουρας, Μ., (2015-2016) *Τεχνητή Νοημοσύνη.* Ανακτήθηκε στις 15/06/2024 από <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>
* Τσούλος, Ι., (χ.χ.) *Τεχνητά Νευρωνικά Δίκτυα.* Ανακτήθηκε στις 08/06/2024 από <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>(Files - Machine Learning - ΔΙΑΛΕΞΕΙΣ - lecture 5)

----------------------------------

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

# ΠΑΡΑΡΤΗΜΑ

*[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*

[Οπισθόφυλλο. Κενή σελίδα]

1. Χρησιμοποιείται στο δίκτυο Adaline (Διαμαντάρας, 2007)· (Δούνιας και Καραμπότσης, χ.χ.)· (Τσούλος, χ.χ) [↑](#footnote-ref-0)
2. Αλλιώς αναφέρεται και ως συνάρτηση κατωφλίου (Λύκας, χ.χ.) ή συνάρτηση Heaviside. (Haykin, 2010) Χρησιμοποιείται στον νευρώνα McCullich & Pitts, αλλά και στο δίκτυο Perceptron (Δούνιας και Καραμπότσης, χ.χ.) Δεν προτιμάται στο 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-1)
3. Η συνάρτηση αυτή συμβολίζεται και ως sgn από το λατινικό ‘signum’ που θα πει πρόσημο (Διαμαντάρας, 2007). [↑](#footnote-ref-2)
4. 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)· (Τσούλος, χ.χ.). [↑](#footnote-ref-3)
5. Η μεταβλητή k αντιπροσωπεύει τον τρέχον αριθμό επανάληψης. Επομένως ο πίνακας βαρών ενημερώνεται σύμφωνα με τιμές τιμές των βαρών που είχε στην προηγούμενη κατάσταση, προσθέτοντας τον ρυθμό μάθησης β πολλαπλασιασμένο με τη διαφορά που προκύπτει μεταξύ της επιθυμητής και της πραγματικής εξόδου του τρέχον προτύπου, πολλαπλασιασμένο με τις τιμές που έχει το ίδιο το πρότυπο. [↑](#footnote-ref-4)
6. Το μέσο τετραγωνικό σφάλμα προκύπτει από τον μέσο όρο των τετραγώνων των διαφορών μεταξύ των επιθυμητών και των πραγματικών τιμών του μοντέλου. [↑](#footnote-ref-5)
7. Προκύπτει από το γινόμενο της παραγώγου της συνάρτησης ενεργοποίησης ως προς τη συνολική είσοδο που λαμβάνει ο νευρώνας με τη διαφορά ανάμεσα στην πραγματική και επιθυμητή έξοδο του νευρώνα. [↑](#footnote-ref-6)
8. Προκύπτει από το γινόμενο της παραγώγου της συνάρτησης ενεργοποίησης ως προς τη συνολική είσοδο που λαμβάνει ο νευρώνας με το άθροισμα των γινομένων των σφαλμάτων του επόμενου στρώματος με τα αντίστοιχα βάρη συνδέσεων. [↑](#footnote-ref-7)
9. Τα βάρη ενημερώνονται αφαιρόντας τον συντελεστή του ρυθμού μάθησης πολλαπλασιασμένο με την παράγωγο της συνάρτησης απώλειας ως προς το βάρος. [↑](#footnote-ref-8)