

### Ασκήσεις μελέτης της 19<sup>ης</sup> διάλεξης

**19.1.** Στο νευρωνικό δίκτυο της διαφάνειας 9, που υλοποιεί την πύλη XOR, υπολογίστε το νέο διάνυσμα χαρακτηριστικών που παράγουν οι δύο νευρώνες του κρυφού επιπέδου (δηλαδή το διάνυσμα που περιέχει τις εξόδους των δύο πυλών AND) για κάθε ένα από τα δυνατά αρχικά διανύσματα χαρακτηριστικών της εισόδου του δικτύου. Δείξτε ότι τα παραγόμενα νέα διανύσματα είναι γραμμικά διαχωρίσιμα (άρα μπορούν να διαχωριστούν από τον νευρώνα του επιπέδου εξόδου).

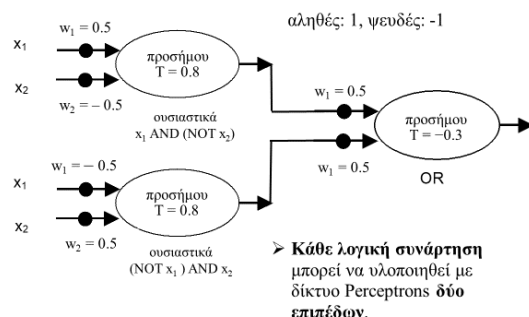
*Απάντηση:* Έστω  $\langle x_1, x_2 \rangle$  το διάνυσμα στην είσοδο του δικτύου. Έστω ότι οι τιμές 1, -1 παριστάνουν τις λογικές τιμές True και False, αντίστοιχα, όπως στη διαφάνεια 10. Έστω  $n_1, n_2$  οι έξοδοι της επάνω και κάτω πύλης AND, αντίστοιχα, της διαφάνειας 10. Τα δύο δυνατά διανύσματα εισόδου για τα οποία η έξοδος του XOR είναι False, δηλαδή τα  $\langle x_1 = -1, x_2 = -1 \rangle$  και  $\langle x_1 = 1, x_2 = 1 \rangle$  απεικονίζονται και τα δύο στο νέο σημείο  $\langle n_1, n_2 \rangle = \langle -1, -1 \rangle$ . Τα δύο δυνατά διανύσματα εισόδου για τα οποία η έξοδος του XOR είναι True, δηλαδή τα  $\langle x_1 = -1, x_2 = 1 \rangle$  και  $\langle x_1 = 1, x_2 = -1 \rangle$  απεικονίζονται στα νέα σημεία  $\langle n_1, n_2 \rangle = \langle -1, 1 \rangle$  και  $\langle n_1, n_2 \rangle = \langle 1, -1 \rangle$  αντίστοιχα. Τα συνολικά τρία νέα σημεία  $\langle n_1, n_2 \rangle$  (ένα με ετικέτα False, δύο με ετικέτα True) είναι τώρα γραμμικά διαχωρίσιμα.

**19.2 (α)** Δείξτε ότι το Perceptron με σιγμοειδή συνάρτηση ενεργοποίησης είναι γραμμικός διαχωριστής.

*Απάντηση:* Για κάθε είσοδο, το Perceptron με σιγμοειδή συνάρτηση ενεργοποίησης παράγει στην έξοδό του την τιμή  $\sigma(\vec{w} \cdot \vec{x}) = \sigma(w_n x_n + \dots + w_1 x_1 + w_0)$ , θεωρώντας ότι έχουμε πάντα  $x_0 = 1$ , όπου  $\vec{w}$  το διάνυσμα βαρών που έχει μάθει το Perceptron,  $\vec{x}$  το διάνυσμα χαρακτηριστικών της εισόδου και  $\sigma$  η σιγμοειδής συνάρτηση. Κατατάσσουμε την είσοδο στη θετική κατηγορία αν  $\sigma(\vec{w} \cdot \vec{x}) \geq 0.5$  και στην αρνητική αν  $\sigma(\vec{w} \cdot \vec{x}) < 0.5$ . Επειδή  $\sigma(z) \geq 0.5$  αν και μόνο αν  $z \geq 0$ , έπεται ότι κατατάσσουμε στη θετική κατηγορία τα  $\vec{x}$  για τα οποία  $\vec{w} \cdot \vec{x} \geq 0$ , δηλαδή αυτά που βρίσκονται είτε πάνω στο υπερ-επίπεδο  $\vec{w} \cdot \vec{x} = 0$  είτε στον θετικό ημιχώρο του υπερ-επιπέδου  $\vec{w} \cdot \vec{x} = 0$  (πιο «πάνω» από το υπερ-επίπεδο), ενώ κατατάσσουμε στην αρνητική κατηγορία κάθε άλλο  $\vec{x}$ . Η διαχωριστική υπερ-επιφάνεια είναι το υπερ-επίπεδο  $\vec{w} \cdot \vec{x} = 0$ , που είναι γραμμικός διαχωριστής.

(β) Δείξτε ότι ένα δίκτυο δύο επιπέδων με Perceptrons (παρόμοιο εκείνου που χρησιμοποιήσαμε για να υλοποιήσουμε την πύλη XOR) μπορεί να μάθει οποιαδήποτε λογική συνάρτηση. Θεωρήστε ότι το αληθές παριστάνεται με 1 και το ψευδές με -1. Θεωρήστε επίσης ότι όλοι οι νευρώνες έχουν συνάρτηση ενεργοποίησης προσήμου, με κατώφλι 0.8 στους νευρώνες του πρώτου επιπέδου. Τα βάρη των εισόδων του μοναδικού νευρώνα του δεύτερου επιπέδου είναι όλα 0.5. Το κατώφλι του νευρώνα του δεύτερου επιπέδου είναι -0.3, αν έχουμε δύο νευρώνες στο πρώτο επίπεδο. Αν έχουμε περισσότερους νευρώνες στο πρώτο επίπεδο, το κατώφλι του νευρώνα του δεύτερου επιπέδου επιλέγεται έτσι, ώστε να αρκεί η ενεργοποίηση ενός μόνο νευρώνα του πρώτου επιπέδου για να ενεργοποιηθεί ο νευρώνας του δεύτερου επιπέδου (π.χ. για δύο νευρώνες στο πρώτο επίπεδο, μπορούμε να επιλέξουμε κατώφλι -0.7 στο δεύτερο επίπεδο).

#### Υλοποίηση XOR με δύο επίπεδα



Απάντηση: Κάθε λογική συνάρτηση μπορεί να οριστεί με έναν πίνακα αληθείας. Για παράδειγμα, ο διπλάνος πίνακας αληθείας ορίζει μια λογική συνάρτηση τριών μεταβλητών ( $X, Y, Z$ ) με απόκριση  $C$ . Βάζουμε στο πρώτο επίπεδο τόσους νευρώνες όσες είναι οι γραμμές του πίνακα αληθείας για τις οποίες  $C = T$  (4 νευρώνες στο παράδειγμά μας). Κάθε νευρώνας του πρώτου επιπέδου παίρνει ως είσοδο όλες τις μεταβλητές ( $X, Y, Z$ ). Φροντίζουμε ο κάθε νευρώνας του πρώτου επιπέδου να λειτουργεί ως πύλη AND και να ενεργοποιείται όταν οι τιμές των μεταβλητών έχουν τις τιμές μιας (διαφορετικής για κάθε νευρώνα) γραμμής του πίνακα αληθείας στην οποία  $C = T$  και μόνο τότε. Στο παράδειγμά μας, ο νευρώνας για την πρώτη γραμμή του πίνακα αληθείας θα μπορούσε να έχει βάρη  $-0.3$  και στις τρεις εισόδους του, ώστε να ενεργοποιείται (έξοδος 1) όταν και οι τρεις εισοδοί του είναι  $F$  ( $-1$ ) και μόνο τότε. Ο νευρώνας του πρώτου επιπέδου για τη δεύτερη γραμμή του πίνακα αληθείας θα είχε βάρη  $-0.3$  και  $-0.3$  για τα  $X, Y$ , ενώ βάρος  $+0.3$  για το  $Z$ , ώστε να πυροδοτείται αν και μόνο αν  $X = -1$  ( $F$ ),  $Y = -1$  ( $F$ ),  $Z = 1$  ( $T$ ). Αντίστοιχα για τους νευρώνες των άλλων δύο γραμμών του πίνακα αληθείας όπου  $C = T$ . Ο μοναδικός νευρώνας του δεύτερου επιπέδου θα λειτουργούσε ως πύλη OR, όπως στην υλοποίηση της πύλης XOR.

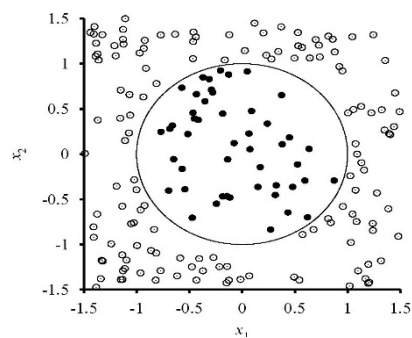
X	Y	Z	C
F	F	F	T
F	F	T	T
F	T	F	T
F	T	T	F
T	F	F	F
T	F	T	T
T	T	F	F
T	T	T	F

(γ) Εξηγήστε γιατί ένα νευρωνικό δίκτυο Perceptrons δύο επιπέδων, σαν αυτό του προηγούμενου σκέλους, μπορεί με τα κατάλληλα βάρη να παραστήσει (άρα δυνητικώς να μάθει) τέλεια κάθε σύνολο δεδομένων εκπαίδευσης, όταν χρησιμοποιούνται μόνο Boolean ιδιότητες και προβλέπεται μία Boolean τιμή, αρκεί να έχει αρκετούς νευρώνες στο πρώτο επίπεδο και να μην υπάρχουν ασυνεπή παραδείγματα εκπαίδευσης, αλλά ενδέχεται να μην έχει καλές επιδόσεις σε δεδομένα αξιολόγησης.

Απάντηση: Τα παραδείγματα εκπαίδευσης μπορούν να παρασταθούν με έναν (πιθανώς ημιτελή) πίνακα αληθείας σαν τον πίνακα στα δεξιά, όπου θεωρούμε ότι έχουμε μόνο 4 παραδείγματα εκπαίδευσης τα οποία προέρχονται από τον πίνακα αληθείας του προηγούμενου σκέλους (εδώ έχουμε μόνο τις πρώτες δύο και τις τελευταίες δύο γραμμές του πίνακα του προηγούμενου σκέλους). Κατασκευάζουμε πάλι ένα δίκτυο Perceptrons δύο επιπέδων, όπως στο προηγούμενο σκέλος. Εδώ το πρώτο επίπεδο θα έχει μόνο δύο νευρώνες (πύλες AND), αφού έχουμε μόνο δύο θετικά ( $C = T$ ) παραδείγματα εκπαίδευσης. Το δίκτυο θα απαντά σωστά όταν οι τιμές των  $X, Y, Z$  αντιστοιχούν σε μια από τις 4 γραμμές του πίνακα των παραδειγμάτων εκπαίδευσης (ουσιαστικά έχει απομνημονεύσει τις γραμμές των δύο παραδειγμάτων με  $C = T$  και απαντά  $C = F$  σε κάθε άλλη περίπτωση), αλλά στις υπόλοιπες γραμμές του πίνακα του προηγούμενου σκέλους, που θα μπορούσαν να χρησιμοποιούνται ως δεδομένα αξιολόγησης, δεν θα συμπεριφέρεται πάντα σωστά. Θα απαντά λανθασμένα  $C = F$  στον τρίτο από την αρχή και στον τρίτο από το τέλος συνδυασμό τιμών  $X, Y, Z$  του πίνακα του προηγούμενου σκέλους, αντί της σωστής απόκρισης  $C = T$ .

X	Y	Z	C
F	F	F	T
F	F	T	T
T	T	F	F
T	T	T	F

**19.3.** (α) Θέλουμε να εκπαιδεύσουμε ένα μεμονωμένο Perceptron, ώστε να διαχωρίζει τις περιπτώσεις των δύο κατηγοριών (μαύρες και άσπρες τελείες, μέσα και έξω από έναν κύκλο) του σχήματος στα δεξιά. Υπάρχουν δύο ιδιότητες, που αντιστοιχούν στους άξονες. Εξηγήστε γιατί αποκλείεται να μάθει το Perceptron να διαχωρίζει σωστά τις δύο κατηγορίες με την υπάρχουσα διανυσματική παράσταση των περιπτώσεων.



Σχήμα από το βιβλίο των Russel και Norvig.

Απάντηση: Το Perceptron είναι γραμμικός διαχωριστής, δηλαδή κατά την εκπαίδευση μαθαίνει ένα σημείο (αν έχουμε μόνο μία ιδιότητα), μια ευθεία γραμμή (αν έχουμε δύο ιδιότητες), ένα επίπεδο (αν έχουμε τρεις ιδιότητες) ή γενικότερα υπερ-επίπεδο (αν έχουμε περισσότερες ιδιότητες) και κατά την αξιολόγηση κατατάσσει τις περιπτώσεις (διανύσματα

ιδιοτήτων) που του δίνουμε σε δύο κατηγορίες, ανάλογα με το αν το διάνυσμα κάθε περίπτωσης βρίσκεται πάνω ή κάτω από το υπερ-επίπεδο. Οι περιπτώσεις του σχήματος δεν είναι γραμμικά διαχωρίσιμες στο διανυσματικό χώρο του σχήματος (δεν υπάρχει ευθεία γραμμή που να διαχωρίζει τις μαύρες από τις άσπρες τελείες). Επομένως, αποκλείεται ένα Perceptron να καταφέρει να μάθει μια ευθεία που να διαχωρίζει πλήρως τις περιπτώσεις των δύο κατηγοριών, αν παριστάνονται όπως στο σχήμα της εκφώνησης.

(β) Να προτείνετε ένα μετασχηματισμό της υπάρχουσας διανυσματικής παράστασης, που να μετατρέπει κάθε περίπτωση (τελεία) σε έναν μόνο πραγματικό αριθμό (να έχουμε μία μόνο ιδιότητα), ώστε να είναι πλέον δυνατόν (με τη νέα παράσταση) να μάθει το Perceptron να διαχωρίζει τις περιπτώσεις των δύο κατηγοριών του σχήματος. Ποιος θα ήταν ο μετασχηματισμός και γιατί θα μπορούσε πλέον το Perceptron να μάθει να διαχωρίζει τις δύο κατηγορίες;

*Απάντηση:* Μπορούμε να παραστήσουμε κάθε περίπτωση (τελεία) με την απόστασή της από το κέντρο του κύκλου. Τότε όλες οι περιπτώσεις διατάσσονται πάνω σε ένα μόνο άξονα, που παριστάνει την απόσταση (τη μία και μοναδική ιδιότητα). Όλες οι μαύρες τελείες θα βρίσκονται αριστερά από την τιμή του άξονα που θα αντιστοιχεί στην ακτίνα του κύκλου (περίπου 1) και όλες οι άσπρες τελείες θα βρίσκονται δεξιά από εκείνη την τιμή. Με την νέα παράσταση, οι περιπτώσεις είναι γραμμικά διαχωρίσιμες και επομένως είναι δυνατόν το Perceptron να μάθει να τις διαχωρίζει.

**19.4.** (α) Δύο φοιτητές συζητούν τη σχέση του Perceptron με τους ταξινομητές λογιστικής παλινδρόμησης. Ο πρώτος ισχυρίζεται ότι ένας ταξινομητής λογιστικής παλινδρόμησης δύο κατηγοριών δεν είναι παρά ένα Perceptron με σιγμοειδή συνάρτηση ενεργοποίησης  $\Phi(S) = 1/(1 + e^{-S})$ . Για να υποστηρίξει την άποψή του, έγραψε τους τύπους που υπολογίζουν, για κάθε διάνυσμα εισόδου  $\vec{x}$ , (i) την έξοδο του Perceptron και (ii) την πιθανότητα με την οποία κατατάσσει ο ταξινομητής λογιστικής παλινδρόμησης το  $\vec{x}$  στη θετική κατηγορία. Γράψτε τους δύο αυτούς τύπους. Τι δείχνουν;

*Απάντηση:* Για κάθε διάνυσμα εισόδου  $\vec{x}$ , η έξοδος ενός Perceptron με την παραπάνω σιγμοειδή συνάρτηση ενεργοποίησης είναι:

$$\Phi\left(\sum w_l x_l\right) = \Phi(\vec{w} \cdot \vec{x}) = 1/(1 + e^{-\vec{w} \cdot \vec{x}})$$

Η πιθανότητα με την οποία κατατάσσει ένας ταξινομητής λογιστικής παλινδρόμησης ένα διάνυσμα εισόδου  $\vec{x}$  στη θετική κατηγορία είναι:

$$P(c_+ | \vec{x}) = 1/(1 + e^{-\vec{w} \cdot \vec{x}})$$

Οι δύο τύποι δείχνουν ότι αν χρησιμοποιήσουμε τα ίδια βάρη  $\vec{w}$ , η έξοδος του Perceptron θα είναι ακριβώς η ίδια με την  $P(c_+ | \vec{x})$  του ταξινομητή λογιστικής παλινδρόμησης, για κάθε  $\vec{x}$ .

(β) Ο δεύτερος φοιτητής ισχυρίζεται ότι ένας ταξινομητής λογιστικής παλινδρόμησης δύο κατηγοριών δεν είναι απλά ένα Perceptron με σιγμοειδή συνάρτηση ενεργοποίησης. Για να υποστηρίξει την άποψή του, υλοποίησε ένα Perceptron με σιγμοειδή συνάρτηση ενεργοποίησης, όπως περιγράφεται στις διαφάνειες του μαθήματος. Υλοποίησε, επίσης, έναν ταξινομητή λογιστικής παλινδρόμησης δύο κατηγοριών (με ανάβαση κλίσης, αλλά και με στοχαστική ανάβαση κλίσης), όπως περιγράφεται στις διαφάνειες. Εκπαίδευσε τις δύο υλοποιήσεις πάνω στα ίδια δεδομένα εκπαίδευσης και έδειξε ότι οι αποφάσεις που παράγουν, σε ίδια δεδομένα αξιολόγησης, είναι πολύ διαφορετικές. Σε τι οφείλονται οι διαφορές στις αποφάσεις των δύο υλοποιήσεων που παρατήρησε ο δεύτερος φοιτητής;

*Απάντηση: Οι διαφορές οφείλονται στο ότι κατά την εκπαίδευση το Perceptron προσπαθεί να ελαχιστοποιήσει (στις διαφάνειες με στοχαστική κατάβαση κλίσης) το συνολικό (τετραγωνικό) σφάλμα στα δεδομένα εκπαίδευσης, ενώ ο ταξινομητής λογιστικής παλινδρόμησης προσπαθεί να μεγιστοποιήσει (με ανάβαση κλίσης ή στοχαστική ανάβαση κλίσης κ.λπ.) τη δεσμευμένη πιθανοφάνεια των παραδειγμάτων εκπαίδευσης. Οι δύο αλγόριθμοι, δηλαδή, χρησιμοποιούν κατά την εκπαίδευση διαφορετική αντικειμενική συνάρτηση. Η διαφορά αυτή οδηγεί και σε διαφορετικούς κανόνες ενημέρωσης βαρών, όπως φαίνεται και στις διαφάνειες.*

(γ) Τελικά ποιος φοιτητής έχει δίκιο και γιατί; Αν κανείς δεν έχει πλήρως δίκιο, ποια είναι κατά τη γνώμη σας η σχέση μεταξύ των δύο αλγορίθμων μάθησης;

*Απάντηση: Σύμφωνα με τα παραπάνω, αν χρησιμοποιήσουμε τα ίδια βάρη  $\bar{w}$ , η έξοδος του Perceptron (με σιγμοειδή συνάρτηση ενεργοποίησης) θα είναι ακριβώς η ίδια με την  $P(c_+ | \bar{x})$  του ταξινομητή λογιστικής παλινδρόμησης, για κάθε  $\bar{x}$ . Επομένως, τόσο το Perceptron όσο και ο ταξινομητής λογιστικής παλινδρόμησης θα παράγουν τα ίδια αποτελέσματα κατά το στάδιο της χρήσης, αν κατά το στάδιο της εκπαίδευσης έχουν μάθει τα ίδια βάρη. Κατά το στάδιο της εκπαίδευσης, όμως, δεν μαθαίνουν εν γένει οπωσδήποτε τα ίδια βάρη, γιατί χρησιμοποιούν διαφορετικές αντικειμενικές συναρτήσεις, όπως προαναφέρθηκε. Επομένως, ο δεύτερος φοιτητής έχει δίκιο, αν και το επιχείρημά του δεν είναι πλήρως πειστικό, γιατί οι διαφορές στις αποφάσεις των δύο υλοποιήσεων θα μπορούσαν να οφείλονται σε διαφορές στις λεπτομέρειες των υλοποιήσεων.*