

Άσκηση

Έστω ότι θέλουμε να κάνουμε ταξινόμηση μήλων και πορτοκαλιών με τη χρήση ενός Perceptron.

- 1) Να σχεδιαστεί (διαγραμματικά) ένα σύστημα αυτόματης ταξινόμησης με χρήση Νευρωνικών Δικτύων.
- 2) Να περιγράψετε πιο θα είναι το διάνυσμα χαρακτηριστικών.
- 3) Χρησιμοποιήστε τον αλγόριθμο του Perceptron για να επιλύσετε αυτό το απλό πρόβλημα ταξινόμησης προτύπων, όπου τα διανύσματα χαρακτηριστικών είναι:

$$\{ \mathbf{x}_1 = [1, -1, -1]^T, t_1 = 0 \text{ και } \mathbf{x}_2 = [1, 1, -1]^T, t_2 = 1 \}$$

Θεωρήστε ότι το διάνυσμα των βαρών έχει αρχική τιμή:

$$\mathbf{w}(0) = [0.5, -1, -0.5]^T$$

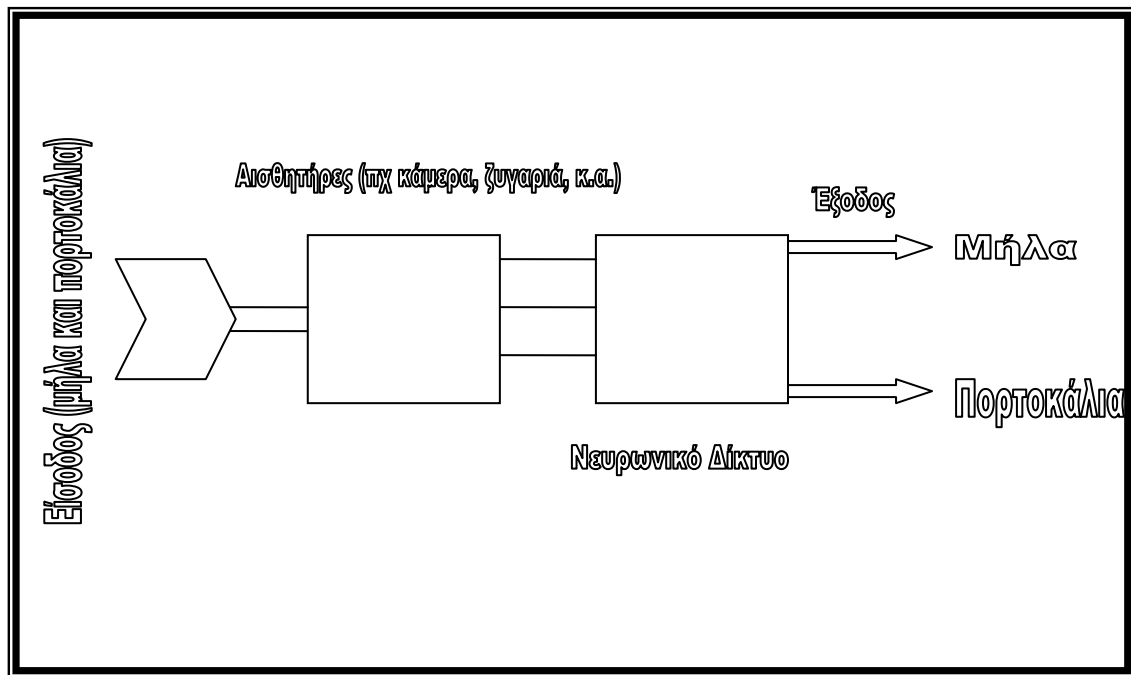
Και η πόλωση αρχικά είναι $b = w_0 = -0.5$, ενώ ο ρυθμός μάθησης η είναι ίσος με 1.

Ως συνάρτηση ενεργοποίησης χρησιμοποιείτε την: $g(u) = \begin{cases} 1, u \geq 0 \\ 0, u < 0 \end{cases}$

Απάντηση

1) Αρχικά χρειαζόμαστε αισθητήρες ώστε να παίρνουμε μετρήσεις από τα φρούτα που περνούν από τη γραμμή παραγωγής προκειμένου να μπορέσουμε να κάνουμε το διαχωρισμό ανάμεσα στα μήλα και στα πορτοκάλια. Οι μετρήσεις από τους αισθητήρες θα μετατρέπονται σε διανύσματα χαρακτηριστικών τα οποία θα τροφοδοτούνται στο Νευρωνικό μας δίκτυο. Ένα σχήμα αυτής της διαδικασίας είναι το σχήμα Ασκ.1.

2) Ο αισθητήρας της εισόδου τροφοδοτεί το Perceptron με τα χαρακτηριστικά των φρούτων (σε ψηφιακή μορφή). Τα χαρακτηριστικά αυτά μπορεί να είναι το χρώμα, το σχήμα και το βάρος (μπορείτε να σκεφτείτε και κάποια άλλα χαρακτηριστικά που θα μπορούσαν να ληφθούν υπόψη;). Το Perceptron στη συνέχεια θα αποφασίζει αν το φρούτο στην είσοδο είναι μήλο ή πορτοκάλι. Ας υποθέσουμε ότι όταν έχουμε έξοδο 0 αποφασίζουμε για τα μήλα και όταν έχουμε έξοδο 1 για τα πορτοκάλια.



Σχήμα Ασκ. Διάγραμμα συστήματος ταξινόμησης μήλων και πορτοκαλιών

3) Από την εκφώνηση γνωρίζουμε ότι τα πρότυπα εκπαίδευσης θα είναι:

$$\mathbf{x}_1 = \begin{bmatrix} +1 \\ -1 \\ -1 \end{bmatrix}, t_1 = 0 \text{ και } \mathbf{x}_2 = \begin{bmatrix} +1 \\ +1 \\ -1 \end{bmatrix}, t_2 = 1$$

Ενώ οι αρχικές τιμές των βαρών είναι: $\mathbf{w}(0) = [0.5, -1, -0.5]^T$ και της πόλωσης: $w_0 = -0.5$

Γνωρίζουμε επίσης ότι σύμφωνα με τον αλγόριθμο του Perceptron ισχύει:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \eta \cdot [t(k) - o(k)] \cdot \mathbf{x}(k)$$

1^{ος} κύκλος (epoch) εκπαίδευσης

Βήμα 1^ο: παρουσιάζουμε το \mathbf{x}_1 στην είσοδο του Perceptron:

$$u(0) = \mathbf{w}(0)^T \cdot \mathbf{x}(0) + w_0(0) \stackrel{\mathbf{x}(0)=\mathbf{x}_1}{=} [0.5, -1, -0.5] \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} - 0.5 = 1.5$$

Επομένως $o(0) = 1$ και $t(0) = t_1 = 0$ οπότε:

$$\mathbf{w}(1) = \mathbf{w}(0) + 1 \cdot [0 - 1] \cdot \mathbf{x}(0) = [0.5, -1, -0.5]^T - 1 \cdot [1, -1, -1]^T = [-0.5, 0, 0.5]^T$$

$$w_0(1) = w_0(0) + 1 \cdot [0 - 1] = -0.5 - 1 = -1.5$$

Βήμα 2^ο: παρουσιάζουμε το x_2 στην είσοδο του δικτύου

$$u(1) = \mathbf{w}(1)^T \cdot \mathbf{x}(1) + w_0(1) \stackrel{\mathbf{x}(1)=\mathbf{x}_2}{=} [-0.5, 0, 0.5] \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - 1.5 = -2.5$$

Επομένως $o(1) = 0$ και $t(1)=t_2=1$ οπότε:

$$\mathbf{w}(2) = \mathbf{w}(1) + 1 \cdot [1 - 0] \cdot \mathbf{x}(1) = [-0.5, 0, 0.5]^T + 1 \cdot [1, 1, -1]^T = [0.5, 1, -0.5]^T$$

$$w_0(2) = w_0(1) + 1 \cdot [1 - 0] = -1.5 + 1 = -0.5$$

2^{ος} κύκλος (epoch) εκπαίδευσης

Βήμα 1^ο: παρουσιάζουμε το x_1 πάλι στην είσοδο του δικτύου

$$u(2) = \mathbf{w}(2)^T \cdot \mathbf{x}(2) + w_0(2) \stackrel{\mathbf{x}(2)=\mathbf{x}_1}{=} [0.5, 1, -0.5] \cdot \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} - 0.5 = -0.5$$

Επομένως $o(2)=0$ και $t(2)=t_1=0$, οπότε δεν θα γίνει καμία διόρθωση στα βάρη και την πόλωση. Δηλαδή:

$$\mathbf{w}(3) = \mathbf{w}(2) = [0.5, 1, -0.5]^T \text{ και } w_0(3) = w_0(2) = -0.5$$

Βήμα 2^ο: παρουσιάζουμε το x_2 πάλι στην είσοδο του δικτύου

$$u(3) = \mathbf{w}(3)^T \cdot \mathbf{x}(3) + w_0(3) \stackrel{\mathbf{x}(3)=\mathbf{x}_2}{=} [0.5, 1, -0.5] \cdot \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - 0.5 = 1.5$$

Επομένως $o(3)=1$ και $t(3)=t_2=1$, οπότε δεν θα γίνει καμία διόρθωση στα βάρη και την πόλωση. Δηλαδή:

$$\mathbf{w}(4) = \mathbf{w}(3) = [0.5, 1, -0.5]^T \text{ και } w_0(4) = w_0(3) = -0.5$$

Άρα αφού τα δείγματα κατηγοριοποιούνται σωστά και κατά συνέπεια δε γίνονται αλλαγές το δίκτυο έχει μάθει και έχει διάνυσμα βαρών:

$$\mathbf{w} = [0.5, 1, -0.5]^T \text{ και } w_0 = -0.5$$

