

# Εκπαίδευση Νευρωνικού Δικτύου με Backpropagation

## Εργασία 5

14 Ιουνίου 2017

Υλοποιήστε τον αλγόριθμο backpropagation όπως παρουσιάζεται στη διάλεξη 11 (*neural-networks*) και εκπαιδεύστε ένα τεχνητό νευρωνικό δίκτυο ώστε να επιλύει το πρόβλημα της κατηγοριοποίησης δεδομένων (*data classification*).

### Για την κατηγοριοποίηση:

- Χρησιμοποιήστε το σετ δεδομένων iris dataset που είναι διαθέσιμο μέσω του MATLAB με την εντολή `[x y] = iris_dataset;` και κατηγοριοποιήστε τα διανύσματα τεσσάρων διαστάσεων σε τρεις κατηγορίες χρησιμοποιώντας τα μισά από τα διαθέσιμα διανύσματα για εκπαίδευση.
- Χρησιμοποιήστε την  $h(x) = \frac{1}{1+e^{-bx}}$  για συνάρτηση ενεργοποίησης των νευρώνων
- Επιλέξτε εσείς τον αριθμό των κρυφών στρώματων, τον αριθμό των νευρώνων του κάθε στρώματος καθώς και τον αριθμό των εποχών που πρέπει να εκπαιδευτεί το δίκτυο.
- Χρησιμοποιήστε τις παραμέτρους  $\lambda = 0.8, b = 1$

### Για την υλοποίηση:

- Ακολουθήστε τον αλγόριθμο στη διαφάνεια 23 της διάλεξης 11 σε συνδυασμό με την περιγραφή του αλγορίθμου με μορφή πινάκων που δίνεται παρακάτω.
- Φροντίστε τα *gradient vectors* να είναι πάντα κανονικοποιημένα.
- Η διόρθωση των βαρών μπορεί να γίνεται για κάθε σημείο ή στο τέλος κάθε εποχής (μπορείτε να δοκιμάσετε και τα δύο, η αλλαγή σε κώδικα που απαιτείται είναι πολύ μικρή).
- Δοκιμάστε την ορθότητα του κώδικά σας αρχικά σε συνθετικά δεδομένα με προκαθορισμένη από εσάς δομή (δείτε τα παραδείγματα του perceptron από τις διαλέξεις)

**Παραδώστε:** Τον κώδικά σας σε MATLAB και μια σύντομη αναφορά (όχι πάνω από τρεις σελίδες) που να φαίνονται οι καμπύλες εκμάθησης του δικτύου και τα ποσοστά επιτυχίας της κατηγοριοποίησης.

**Καταληκτική ημερομηνία:** Παρασκευή 30 Ιουνίου 2017.

### Αλγόριθμος σε μορφή πινάκων:

- *Forward path*:  $i = 1 : L$  όπου  $L$  ο αριθμός των στρωμάτων,  $it = 1 : E$ , όπου  $E$  ο αριθμός των εποχών εκπαίδευσης.

$$v^{[i]} = W^{[i]} u^{[i]}, u^{[i+1]} = h(v^{[i]})$$

$$e = d - u^L, J_t = e^T e, J(it) = J(it) + J_t$$

- *Backward path*:

$$\delta^{[i]} = (d - y)h'(v^{[i]}), i = L$$

$$\delta^{[i]} = h'(x)(W^{[i+1]})^T \delta^{[i+1]}, i = 1 : L - 1$$

$$\frac{\partial J_t}{\partial W^{[i]}} = -\delta^{[i]} (u^{[i]})^T, i = 1 : L$$

$$(\frac{\partial J}{\partial W^{[i]}})_{it} = (\frac{\delta J}{\partial W})_{it} + \frac{\partial J_t}{\partial W}$$

- Συμβολισμοί:  $\delta$ : γενικευμένο σφάλμα,  $u$ : γραμμική έξοδος ενός επιπέδου,  $v$ : μη γραμμική έξοδος, είσοδος στο επόμενο επίπεδο,  $\frac{\partial J_t}{\partial W}$ : *gradient*,  $(\frac{\partial J}{\partial W})_{it}$ : συνολικό *gradient* για κάθε εποχή,  $J(it)$ : Μέσο τετραγωνικό σφάλμα για κάθε εποχή.