



Πανεπιστήμιο Κύπρου
Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών
ΗΜΥ 421 Ευφυή Συστήματα
Διδάσκων: Δρ. Στέλιος Τιμοθέου

2^η Σειρά Ασκήσεων

Ημερομηνία παράδοσης: 08/11/2022

- I. Έστω ότι τα \mathcal{X}_0 και \mathcal{X}_1 είναι δύο σύνολα σημείων στον Ευκλείδειο χώρο N διαστάσεων. Τότε τα \mathcal{X}_0 και \mathcal{X}_1 είναι γραμμικά διαχωρίσιμα αν υπάρχουν N πραγματικοί αριθμοί w_1, \dots, w_N ώστε για κάθε σημείο $\mathbf{x} \in \mathcal{X}_0$ να ισχύει $\sum_{i=1}^N w_i x_i < -\epsilon$ και για κάθε σημείο $\mathbf{y} \in \mathcal{X}_1$ να ισχύει $\sum_{i=1}^N w_i y_i > \epsilon$. Θεωρώντας ότι τα σύνολα σημείων \mathcal{X}_0 και \mathcal{X}_1 στο αρχείο `dataLinearSeparability.mat` είναι γραμμικά διαχωρίσιμα για $\epsilon = 10^{-3}$, να βρείτε ένα διάνυσμα w_1, \dots, w_N που πετυχαίνει αυτό το διαχωρισμό.
- Να μοντελοποιήσετε και να επιλύσετε το πρόβλημα με χρήση γραμμικού προγραμματισμού. **(15 μονάδες)**
 - Να επιλύσετε το πρόβλημα με υλοποίηση του αλγορίθμου ταξινόμησης δύο γραμμικών κλάσεων δεδομένων με βάση το μοντέλο Rosenblatt's perceptron. **(10 μονάδες)**
- II. Να αναπτύξετε ένα εξομοιωτή για τεχνητά νευρωνικά δίκτυα (ΤΝΔ) στον οποίο να υλοποιείτε το βασικός αλγόριθμος backpropagation για πολύ-επίπεδα δίκτυα πρόσθια τροφοδότησης (multilayer feedforward neural networks). **(30 μονάδες)**
Ο εξομοιωτής θα πρέπει να υποστηρίζει τις ακόλουθες λειτουργίες:
- Δυνατότητα επιλογής των πιο κάτω παραμέτρων που σχετίζονται με τον εξομοιωτή:
 - αριθμός κρυφών επιπέδων και νευρώνων σε κάθε επίπεδο, **(3 μονάδες)**
 - σταθερά μάθησης (learning constant), **(3 μονάδες)**
 - είδος συνάρτησης ενεργοποίησης για κάθε νευρώνα (γραμμική, σιγμοειδής (sigmoid) και υπερβολική εφαπτομένη (hyperbolic tangent)). **(4 μονάδες)**
 - Κατάλληλη κανονικοποίηση (normalization) των παραδειγμάτων εκπαίδευσης (training examples) ώστε ο μέσος όρος και η διακύμανση της κάθε εισόδου να είναι 0 και 1 αντίστοιχα. **(5 μονάδες)**
 - Δυνατότητα επιλογής συγκεκριμένου ποσοστού από τα διαθέσιμα δεδομένα για την μάθηση (training), την επικύρωση (validation) και τον έλεγχο (testing). **(5 μονάδες)**
 - Δυνατότητα παρακολούθησης της επίδοσης ενός εκπαιδευόμενου ΤΝΔ μέσω της επίδειξης του κόστους εκπαίδευσης $E_{av}(\mathbf{W})$ για τα training and validation sets ως συνάρτηση του αριθμού επαναλήψεων. **(5 μονάδες)**
- III. Χρησιμοποιώντας τον εξομοιωτή, να μελετηθεί η δυνατότητα μάθησης του αναπτυγμένου εξομοιωτή στα ακόλουθα προβλήματα:
- Ταξινόμηση δύο τάξεων για συνθετικά δεδομένα (διαχωρίσιμες τάξεις). **(Βαρύτητα 13%)**
 - Ταξινόμηση δύο τάξεων για συνθετικά δεδομένα με θόρυβο (μη διαχωρίσιμες τάξεις). **(Βαρύτητα 7%)**
- Συγκεκριμένα για κάθε ένα από τα πιο πάνω προβλήματα να δείξετε:

- a) το ακριβές μοντέλο ΤΝΔ που χρησιμοποιήσατε και όλες τις τιμές των παραμέτρων,
- b) την γραφική παράσταση του κόστους εκπαίδευσης $E_{av}(\mathbf{W})$ ως συνάρτηση του αριθμού επαναλήψεων,
- c) το Confusion Matrix καθώς και τις τιμές των μετρικών Accuracy, Precision, Sensitivity και F1-score για τα δεδομένα εκπαίδευσης (training), επικύρωσης (validation) και ελέγχου (testing).

Σημείωση: Τα δεδομένα εκπαίδευσης και ελέγχου (*training and testing sets*) για το πρόβλημα *III.i*, $i=\{1,2\}$, έχουν ονόματα *trainingSet*i*.dat* και *testingSet*i*.dat* αντίστοιχα.