

3^Η ΥΠΟΧΡΕΩΤΙΚΗ ΕΡΓΑΣΙΑ ΣΤΟ ΜΑΘΗΜΑ «ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ – ΒΑΘΙΑ ΜΑΘΗΣΗ»

Να γραφεί πρόγραμμα σε οποιαδήποτε γλώσσα προγραμματισμού το οποίο να υλοποιεί ένα **Radial Basis Function Neural Network** που θα εκπαιδευτεί για να επιλύει ένα από τα προβλήματα που επιλύσατε στις προηγούμενες εργασίες σας ή αντίστοιχο πρόβλημα προσέγγισης συνάρτησης. Ενδεικτικά αναφέρονται τα παρακάτω:

1. Αναγνώριση ψηφίων ή μονών και ζυγών αριθμών στα δεκαδικά ψηφία (0,1,...,9) της MNIST: <http://www.cs.toronto.edu/~roweis/data.html>
2. Διαχωρισμό 2 ή όλων των κλάσεων που υπάρχουν στις Cifar-10 ή SVHN και βρίσκονται στις παρακάτω διευθύνσεις:
<https://www.cs.toronto.edu/~kriz/cifar.html>
<http://ufldl.stanford.edu/housenumbers/>
3. ή να επιλύει οποιοδήποτε πρόβλημα κατηγοριοποίησης πολλών κλάσεων ή προσέγγισης συνάρτησης από τις βάσεις που βρίσκονται στις παρακάτω σελίδες:
<http://archive.ics.uci.edu/ml/>
<http://www.cs.toronto.edu/~roweis/data.html>
<http://www.cs.cmu.edu/~cil/v-images.html>
<https://www.kaggle.com/datasets>

Όπου δεν υπάρχει σύνολο ελέγχου χωρίζεται η βάση τυχαία σε σύνολο εκπαίδευσης (60%) και ελέγχου (40%) ή ακολουθείται τεχνική cross-validation.

Εξαγωγή Χαρακτηριστικών

Για το διαχωρισμό των δειγμάτων μπορεί να μειώνεται πρώτα η διάσταση των δεδομένων χρησιμοποιώντας PCA ώστε να κρατήσετε περισσότερο από 90% της πληροφορίας.

Έκθεση αποτελεσμάτων

Θα πρέπει να γραφεί έκθεση στην οποία να περιγράφονται: ο αλγόριθμος, να δίνονται χαρακτηριστικά παραδείγματα ορθής και εσφαλμένης κατηγοριοποίησης καθώς και ποσοστά επιτυχίας στα στάδια της εκπαίδευσης (training) και του ελέγχου (testing), χρόνος εκπαίδευσης και ποσοστά επιτυχίας για διαφορετικούς αριθμούς κρυφών νευρώνων, τρόπο εκπαίδευσης (K-μέσους, τυχαία επιλογή κέντρων, κτλ) καθώς και διαφορετικές τιμές των παραμέτρων εκπαίδευσης. Να συγκριθεί η απόδοση του RBF σε σχέση με την κατηγοριοποίηση πλησιέστερου γείτονα (Nearest Neighbor) και πλησιέστερου κέντρου κλάσης (Nearest Class Centroid). Να σχολιασθούν τα αποτελέσματα και ο κώδικας.

ΗΜΕΡΟΜΗΝΙΑ ΠΑΡΑΔΟΣΗΣ : 10^η Ιανουαρίου 2021

Για κάθε ημέρα αργοπορημένης υποβολής της εργασίας και για 5 ημέρες μειώνεται η βαθμολογία κατά 10%.

ΗΜΕΡΟΜΗΝΙΑ ΠΑΡΟΥΣΙΑΣΗΣ/ΕΞΕΤΑΣΗΣ : θα οριστεί

Μετά από την παράδοση της εργασίας θα ακολουθήσει προφορική παρουσίαση των τριών εργασιών από τον κάθε φοιτητή. Η παρουσίαση θα αποτελείται από 8-10 Διαφάνειες σε powerpoint:

- 1^η : Στοιχεία Φοιτητή, βάσεις δεδομένων που χρησιμοποιήθηκαν, ιδιαιτερότητες αν υπάρχουν.
- 2^η -9^η : αποτελέσματα για κάθε μία από τις εργασίες, απόδοση, τιμές παραμέτρων, χρόνος, κτλ.
- 10^η : Συνολικά συμπεράσματα - Σχόλια.

Οι διαφάνειες θα πρέπει να υποβληθούν στο elearning το αργότερο μέχρι την προηγούμενη ημέρα από την παρουσίαση/εξέταση.