

Εργασία για το μάθημα Υπολογιστικής Νοημοσύνης 1 για το ακαδημαϊκό έτος 2012-2013

Περιγραφή

Αναγνώριση Χαρακτήρων: Ο σκοπός της οπτικής αναγνώρισης χαρακτήρων είναι να μετατρέψει σαρωμένες εικόνες γραπτού κειμένου σε κείμενο ASCII που είναι δυνατόν να επεξεργαστεί περαιτέρω. Τέτοια συστήματα αναγνώρισης χαρακτήρων περιέχουν διαφορετικά υποσυστήματα:

1. Ένα υποσύστημα για κατάτμηση της σελίδας έτσι ώστε να ταυτοποιηθούν περιοχές που περιέχουν κείμενο και να διαχωριστούν από περιοχές που περιέχουν εικόνες
2. Ένα υποσύστημα για αφαίρεση θορύβου και κατάτμηση χαρακτήρων (ταυτοποίηση υποπεριοχών που αντιστοιχούν σε έναν χαρακτήρα)
3. Ένα υποσύστημα για συσχέτιση ενός γραφικού προτύπου με έναν ASCII κωδικό που αντιστοιχεί σε έναν χαρακτήρα
4. Ένα υποσύστημα για ανακατασκευή του κειμένου τοποθετώντας τους αναγνωρισμένους χαρακτήρες στη θέση τους

Σκοπός της παρούσας εργασίας: Σκοπός της παρούσας εργασίας είναι να σχεδιαστεί και να υλοποιηθεί ένα υποσύστημα (δηλαδή το υποσύστημα 3) για συσχέτιση ενός γραφικού προτύπου με έναν ASCII χαρακτήρα με χρήση τεχνητών νευρωνικών δικτύων. Για λόγους απλότητας, η συγκεκριμένη εργασία θα εστιάσει στην αναγνώριση των δέκα πρώτων αγγλικών κεφαλαίων χαρακτήρων. Συγκεκριμένα θα χρησιμοποιηθεί ένα δημόσια διαθέσιμο σύνολο δεδομένων:

(<http://archive.ics.uci.edu/ml/datasets/Optical+Recognition+of+Handwritten+Digits>)

που περιέχει μετρήσεις για 5620 χειρόγραφους χαρακτήρες.

Τα δεδομένα αυτά προέρχονται από ένα σύνολο 43 ανθρώπων. Τα χαρακτηριστικά που θα χρησιμοποιηθούν σαν είσοδο για τους ταξινομητές της εργασίας προέρχονται από την bitmap 32X32 αναπαράσταση των γραπτών μορφών των χαρακτήρων. Για να μειωθεί η διαστασιμότητα του προβλήματος κάθε χάρτης 32X32 χωρίστηκε σε μη επικαλυπτόμενα κομμάτια 4X4 και ο αριθμός των πίξελ σε κάθε τέτοιο κομμάτι μετρήθηκε. Έτσι, δημιουργήθηκε ένας πίνακας εισόδου 8X8 (δηλαδή μήκους 64) όπου κάθε στοιχείο του λαμβάνει τιμές από 0 έως 16.

Τα δεδομένα χωρίζονται σε σύνολο εκπαίδευσης (optdigits.tra) και σύνολο ελέγχου (optdigits.tes) με τις πρώτες 64 στήλες να είναι οι είσοδοι και την τελευταία στήλη να είναι ο κωδικός αριθμός της επιθυμητής εξόδου (τιμές από 0 έως 9 για τους δέκα χαρακτήρες που προσπαθούμε να αναγνωρίσουμε).

Ερωτήματα

Ερώτημα 1. (10%) Προεπεξεργασία δεδομένων

1.1 (2,5%) Εισάγετε τα δεδομένα στο MATLAB με σκοπό να δημιουργήσετε έναν πίνακα με τα δεδομένα εκπαίδευσης και έναν με τα δεδομένα ελέγχου.

1.2 (2,5%) Κανονικοποιήστε τις αριθμητικές τιμές των χαρακτηριστικών εισόδου του συνόλου δεδομένων σας στο $[-1 \ 1]$.

1.3 (2,5%) Δημιουργήστε το διάνυσμα εξόδου σας με χρήση της κωδικής τιμής εξόδου από 0-9 που παρέχεται για κάθε δείγμα.

Π.χ.

0 -> 0000000001

1 -> 0000000010

2-> 0000000100

κ.ο.κ.

1.4 (2,5%) Αναδιατάξτε τις εγγραφές κάθε συνόλου με τυχαίο τρόπο αφού αρχικοποιήσετε την γεννήτρια του matlab με seed ίσο με το άθροισμα των αριθμών μητρώου των μελών της κάθε ομάδας.

Ερώτημα 2. (50%) Σχεδιασμός και Υλοποίηση Τεχνητού Νευρωνικού Δικτύου

Χρησιμοποιείστε τις έτοιμες συναρτήσεις του MATLAB για να δημιουργήσετε ένα πολυεπίπεδο εμπρός τροφοδότησης νευρωνικό δίκτυο που να λύνει το παραπάνω πρόβλημα. Το νευρωνικό σας δίκτυο θα έχει ένα κρυφό επίπεδο και τόσους κρυφούς νευρώνες όσους προκύπτουν από τον παρακάτω κανόνα:

$$H \approx (I+O)/2$$

όπου H είναι ο αριθμός των νευρώνων του κρυφού επιπέδου, I ο αριθμός εισόδων (64) και O ο αριθμός εξόδων του TN- αντίστοιχα. Οι εξοδοί του τεχνητού νευρωνικού δικτύου θα είναι 10.

Από τις διάφορες συναρτήσεις εκπαίδευσης που παρέχει το matlab χρησιμοποιήστε την «traingdm». Η συγκεκριμένη συνάρτηση εκπαίδευσης υλοποιεί τον αλγόριθμο της πίσω διάδοσης του σφάλματος με χρήση της παραμέτρου ορμής.

Ερώτημα 3. (40%) Πειραματικά Αποτελέσματα και Συμπεράσματα

3.1 (30%) Συμπληρώστε τον παρακάτω πίνακα εκτελώντας από 5 πειράματα για κάθε σύνολο παραμέτρων και παρουσιάζοντας τον μέσο όρο των αποδόσεων. Σαν συνάρτηση ενεργοποίησης χρησιμοποιήστε την λογιστική σε όλους τους νευρώνες του επιπέδου εξόδου και του κρυφού επιπέδου. Σαν κριτήριο τερματισμού της εκπαίδευσης θα χρησιμοποιηθεί ο συνδυασμός των παρακάτω κριτηρίων (έστω και ένα κριτήριο να ισχύει η εκπαίδευση του νευρωνικού δικτύου θα τερματίζει)

- Μέγιστος αριθμός εποχών εκπαίδευσης = 100
- Μέγιστος επιτρεπτός αριθμός αποτυχημένων ελέγχων επικύρωσης (δηλαδή αριθμός διαδοχικών εποχών που το σφάλμα επικύρωσης αυξάνεται. Το σύνολο επικύρωσης στο MATLAB δημιουργείται (με την προεπιλεγμένη του τιμή) λαμβάνοντας το 20% του συνόλου εκπαίδευσης) = 10
- Ελάχιστη επιτρεπόμενη τιμή της κλίσης της συνάρτησης σφάλματος (minimum gradient)= $1e-10$

Για τον υπολογισμό της ακρίβειας πρόβλεψης θεωρείστε ότι το νευρωνικό δίκτυο ταξινομεί στον χαρακτήρα i (από 0-9) ανάλογα με το ποια από τις 10 εξόδους του έχει την μεγαλύτερη τιμή. Για παράδειγμα αν για ένα δείγμα έχουμε την ακόλουθη έξοδο

Νευρώνας εξόδου 0: 0.1
 Νευρώνας εξόδου 1: 0.5
 Νευρώνας εξόδου 2: 0.7
 Νευρώνας εξόδου 3: 0.6
 Νευρώνας εξόδου 4: 0.2
 Νευρώνας εξόδου 5: 0.1
 Νευρώνας εξόδου 6: 0.15
 Νευρώνας εξόδου 7: 0.3
 Νευρώνας εξόδου 8: 0.6
 Νευρώνας εξόδου 9: 0.1

Τότε αυτό το δείγμα θα ταξινομηθεί στον χαρακτήρα 2 (χαρακτήρας 'C') αφού η μέγιστη έξοδος είναι αυτή του νευρώνα 2

A/A	Παράμετρος Μάθησης	Παράμετρος Ορμής	Σύνολο Εποχών Εκπαίδευσης	Μέσο Τετραγωνικό Σφάλμα (σύνολο εκπαίδευσης)	Μέσο Τετραγωνικό Σφάλμα (σύνολο ελέγχου)	Ακρίβεια Πρόβλεψης (σύνολο εκπαίδευσης)	Ακρίβεια Πρόβλεψης Σύνολο Ελέγχου
1	0.01	0.9					
2	0.01	0.6					
3	0.01	0.1					
4	0.1	0.9					
5	0.1	0.6					
6	0.1	0.1					

Σχολιάστε τα αποτελέσματά σας σε σχέση με

- την διαφοροποίηση της απόδοσης με χρήση διαφορετικών τιμών παραμέτρων
- την ικανότητα γενίκευσης
- την ταχύτητα σύγκλισης (ενδεικτικά παραθέστε και κάποιες γραφικές παραστάσεις όπου κρίνετε απαραίτητο)

3.2 (10%) Χρησιμοποιείτε τον διπλάσιο αριθμό κρυφών νευρώνων σε σχέση με το προηγούμενο ερώτημα και επαναλάβετε τα πειράματά σας για το βέλτιστο σύνολο παραμέτρων που εντοπίσατε στο προηγούμενο ερώτημα. Τι παρατηρείτε;

Ερώτημα 4 (Ερώτημα Bonus). Σχεδιασμός και Υλοποίηση Διαπεφής Χρήστη

Δημιουργήστε μια διεπαφή Matlab GUI μέσα από την οποία θα μπορεί κάποιος χρήστης να εκτελέσει όλα τα βήματα της προεπεξεργασίας των δεδομένων, της εκπαίδευσης των τεχνητών νευρωνικών δικτύων και της προβολής των αποτελεσμάτων.

Ερώτημα 5. Απαλλακτική Εργασία

Σχεδιάστε και υλοποιήστε σε γλώσσα της αρεσκείας σας τα υπόλοιπα τρία υπο-συστήματα (δηλαδή τα 1,2 και 4) ενός συστήματος αναγνώρισης χαρακτήρων όπως αυτά αναφέρονται στην εισαγωγή της εργασίας. Συνδέστε τα επιμέρους υποσυστήματα σε ένα ενιαίο σύστημα αναγνώρισης χαρακτήρων.

Παραδοτέα

- Η αναφορά που θα παραδώσετε θα πρέπει να περιέχει εκτενή σχολιασμό των πειραμάτων σας καθώς και πλήρη καταγραφή των αποτελεσμάτων σας.
- Η άσκηση θα πρέπει να υλοποιηθεί σε MATLAB, όπου υπάρχουν και οι κατάλληλες συναρτήσεις για την υλοποίηση των ερωτημάτων.

Σημειώσεις-Σχόλια

- Η ΠΑΡΑΔΟΣΗ ΤΗΣ ΕΡΓΑΣΙΑΣ ΕΠΙΤΡΕΠΕΤΑΙ ΜΕΧΡΙ ΤΗΝ ΗΜΕΡΟΜΗΝΙΑ ΤΗΣ ΓΡΑΠΤΗΣ ΕΞΕΤΑΣΗΣ ΤΟΥ ΜΑΘΗΜΑΤΟΣ.
- Η εργασία είναι αυστηρά των δύο ατόμων. Όσοι δεν έχετε κάποιον να κάνετε την εργασία μαζί παρακαλώ να επικοινωνήσετε με email μαζί μας. Τα ερωτήματα 4 και 5 είναι ατομικά και δεν επιτρέπεται συνεργασία.
- Δειγματοληπτικά θα κληθείτε να εξεταστείτε προφορικά για τις εργασίες σας.
- Μαζί με την αναφορά πρέπει να παραδώσετε και τον κώδικα που υλοποιήσατε στα πλαίσια της εργασίας.
- Για την υλοποίηση του κώδικα μπορείτε να χρησιμοποιήσετε έτοιμες συναρτήσεις λογισμικού που υλοποιούν και εκπαιδεύουν Νευρωνικά Δίκτυα οι οποίες υπάρχουν στο Matlab .
- Η αναφορά και ο κώδικας παραδίδονται σε έντυπη και σε ηλεκτρονική μορφή. Η εργασία θα αποτελέσει το 30% του τελικού βαθμού. Σε περίπτωση που υλοποιήσετε σωστά το ερώτημα 4 η εργασία θα μετρήσει 40%. Σε περίπτωση που υλοποιήσετε το ερώτημα 5 η εργασία θα αποτελέσει το 100% του τελικού σας βαθμού.
- Για οποιαδήποτε διευκρίνιση / ερώτηση μπορείτε να επικοινωνήσετε με τους (ώρες γραφείου 10:00-14:00):
 - Κορφιάτη Αίγλη (korfiati@ceid.upatras.gr)
 - Θεοφιλάτο Κωνσταντίνο (theofilk@ceid.upatras.gr)