



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΕΞΕΡΓΑΣΙΑΣ ΕΙΚΟΝΑΣ ΒΙΝΤΕΟ ΚΑΙ ΠΟΛΥΜΕΣΩΝ

Νευρωνικά Δίκτυα και Ευφυή Υπολογιστικά Συστήματα
2007-2008

**1η Εργαστηριακή Άσκηση: “Μελέτη των Πολυεπίπεδων
Perceptrons και εφαρμογή σε πρόβλημα ταξινόμησης
εικόνας με βάση το θεματικό της περιεχόμενο”**

Εισαγωγή

Τα πολυεπίπεδα δίκτυα εμπρόσθιας τροφοδότησης (Multilayer Perceptrons) αποτελούν μια σημαντική κατηγορία νευρωνικών δικτύων. Συνήθως, ένα τέτοιο δίκτυο αποτελείται από ένα σύνολο αισθητήρων που αποτελούν το επίπεδο εισόδου, ένα ή περισσότερα κρυμμένα επίπεδα στα οποία γίνονται οι διάφοροι υπολογισμοί και ένα επίπεδο εξόδου. Το σήμα που μπαίνει στο επίπεδο εισόδου διαδίδεται κατά την εμπρόσθια διεύθυνση, επίπεδο προς επίπεδο. Ένας πολύ συνηθισμένος αλγόριθμος που εφαρμόζεται για την εκπαίδευσή του είναι ο αλγόριθμος οπίσθιας διάδοσης (Back Propagation).

Τα χαρακτηριστικά ενός του πολυεπίπεδου perceptron είναι τα εξής:

- Κάθε νευρώνας του δικτύου περιλαμβάνει μία μη γραμμική (διαφοριστική) συνάρτηση ενεργοποίησης.
- Το δίκτυο αποτελείται από ένα ή περισσότερα κρυμμένα επίπεδα νευρώνων, τα οποία δεν αποτελούν είσοδο ή έξοδο και του παρέχουν τη δυνατότητα μάθησης.
- Το δίκτυο έχει μεγάλο βαθμό συνεκτικότητας, που καθορίζεται από τις συνάψεις του.

Ο συνδυασμός των παραπάνω χαρακτηριστικών, μαζί με την ικανότητα της μάθησης μέσω της εκπαίδευσης προσδίδει στο δίκτυο μεγάλη υπολογιστική ισχύ. Στη σχεδίαση ενός νευρωνικού δικτύου υπεισέρχονται πολλά προβλήματα που πρέπει να αντιμετωπιστούν. Για πολλά από αυτά δεν υπάρχει συγκεκριμένη λύση. Αντίθετα ο χρήστης πρέπει να τα αντιμετωπίσει στηριζόμενος στην προσωπική του εμπειρία.

Ο σκοπός της άσκησης αυτής είναι η κατασκευή και η μελέτη ενός Multilayer Perceptron (MLP), που θα εκπαιδευτεί με τον κλασικό αλγόριθμο Back Propagation, καθώς και η επίλυση δύο προβλημάτων. Στο πρώτο μέρος της άσκησης, το νευρωνικό δίκτυο θα χρησιμοποιηθεί για την εκτίμηση της τιμής ενός σπιτιού, όταν δίνονται τα υπόλοιπα χαρακτηριστικά του. Στο δεύτερο μέρος της άσκησης, το νευρωνικό δίκτυο θα χρησιμοποιηθεί σε πρόβλημα θεματικής ταξινόμησης εικόνων, με βάση τα οπτικά χαρακτηριστικά τους.

1 Μελέτη ενός Νευρωνικού Δικτύου και επίλυση προβλήματος εκτίμησης τιμής

Το σύνολο δεδομένων που θα χρησιμοποιηθεί για την εκπαίδευση του δικτύου περιλαμβάνει δεδομένα τα οποία αντιστοιχούν σε διάφορα χαρακτηριστικά ενός σπιτιού. Για παράδειγμα, δίνεται η εγκληματικότητα της περιοχής, ο αριθμός των δωματίων του κλπ. Το σύνολο εκμάθησης αποτελείται από 466 παραδείγματα και το σύνολο ελέγχου από 50 παραδείγματα. Σκοπός είναι η εκτίμηση του σπιτιού, με βάση τα υπόλοιπα χαρακτηριστικά του.

1.1 Προεπεξεργασία των δεδομένων

1. Να φορτώσετε τα δεδομένα στο workspace της MATLAB και να φτιάξετε έναν πίνακα που να περιέχει τα χαρακτηριστικά των παραδειγμάτων του συνόλου εκπαίδευσης και έναν που να περιέχει τις επιθυμητές αποκρίσεις του δικτύου για αυτά. Να κάνετε το ίδιο και για το σύνολο ελέγχου.
2. Να εφαρμόσετε Ανάλυση Πρωτογενών Συνιστωσών (Principal Component Analysis - PCA) προκειμένου να αποσυσχετίσετε τα διάφορα χαρακτηριστικά και να μειώσετε τη διάστασή τους, σύμφωνα με τα αποτελέσματα της ανάλυσης. Μπορείτε να εξηγήσετε ποιοτικά ποια χαρακτηριστικά αφαιρέσατε και γιατί;
3. Να επεξεργασθείτε κατάλληλα τα δεδομένα ώστε να έχουν μέση τιμή ίση με το μηδέν και τυπική απόκλιση ίση με ένα. Στη συνέχεια, να κανονικοποιήσετε τα δεδομένα επιλέγοντας κατάλληλο διάστημα. Προσπάθηστε να εξηγήσετε γιατί αυτό είναι επιθυμητό (π.χ. τι θα γινόταν αν όλα τα χαρακτηριστικά είχαν μόνο θετικές ή μόνο αρνητικές τιμές);
4. Να επιλέξετε κατάλληλη συνάρτηση ενεργοποίησης με βάση τον τρόπο με τον οποίο κανονικοποιήσατε τα δεδομένα και να εξηγήσετε την επιλογή σας.
5. * Με βάση τη συνάρτηση ενεργοποίησης που επιλέξατε στο προηγούμενο ερώτημα, υπάρχουν κάποια παραδείγματα στα σύνολα εκμάθησης και εκπαίδευσης που θα δημιουργούν πρόβλημα στον αλγόριθμο back propagation. Να τα εντοπίσετε, να τα αντιμετωπίσετε κατάλληλα και να εξηγήσετε την επιλογή σας.

1.2 Μελέτη της Αρχιτεκτονικής του Νευρωνικού Δικτύου

1. Να κατασκευάσετε ένα Νευρωνικό Δίκτυο MLP και να μελετήσετε την επίδραση του αριθμού των νευρώνων στην απόδοσή του. Να ξεκινήσετε με 1 κρυμμένο επίπεδο και 2 νευρώνες και στη συνέχεια αυξήστε τον αριθμό τους, μέχρι να πετύχετε το καλύτερο δυνατό αποτέλεσμα. Να πειραματιστείτε με την προσθήκη δεύτερου κρυμμένου επιπέδου. Να συγκρίνετε τα αποτελέσματα και να προσπαθήσετε να αιτιολογήσετε τις διαφορές που τυχόν παρατηρήσατε.
2. Δοκιμάστε να εκπαιδεύσετε το δίκτυο με διάφορες τιμές του ρυθμού μάθησης. Τι παρατηρείτε όταν ο ρυθμός μάθησης είναι πολύ μικρός και τι όταν είναι πολύ μεγάλος; Να βρείτε μια τιμή του ρυθμού μάθησης έτσι ώστε το δίκτυο να συνδυάζει ταχύτητα εκπαίδευσης και γενίκευση. Δοκιμάστε να χρησιμοποιήσετε προσαρμοζόμενο ρυθμό μάθησης. Να σχολιάσετε τις διαφορές που παρατηρήσατε στα αποτελέσματα, καθώς και στις καμπύλες μάθησης.
3. Στη συνέχεια να εφαρμόσετε τη μέθοδο ανανέωσης βαρών με τη χρήση ενός όρου ορμής (momentum). Τι διαφορά έχει στην ταχύτητα σύγκλισης και πώς αιτιολογείται;
4. Να εκπαιδεύσετε κατάλληλα το δίκτυο έτσι ώστε να σταματά “πρόωρα” την εκπαίδευση, όταν σε ένα σύνολο επιβεβαίωσης (validation set) το σφάλμα αρχίσει να αυξάνεται για συνεχόμενες εποχές. (Η μέθοδος αυτή ονομάζεται “Early Stopping of Training”). Να επιλέξετε κατάλληλα το σύνολο επιβεβαίωσης, σαν υποσύνολο του συνόλου εκμάθησης. Να εξηγήσετε τι επιτυγχάνει η μέθοδος αυτή με τη βοήθεια κατάλληλου διαγράμματος.
5. Να μελετήσετε την ικανότητα του δικτύου για γενίκευση τροποποιώντας τον αριθμό των εποχών. Τι παρατηρείτε όταν ο αριθμός των εποχών είναι πολύ μικρός και τι όταν είναι πολύ μεγάλος;
6. Με βάση τις παρατηρήσεις σας στα παραπάνω ερωτήματα, να επιλέξετε τις κατάλληλες παραμέτρους για να κατασκευάσετε ένα δίκτυο που να συνδυάζει απλότητα με αποτελεσματικότητα και μεγάλη ικανότητα γενίκευσης.
7. * Υπάρχουν κάποια χαρακτηριστικά τα οποία παίζουν μεγαλύτερο ρόλο στην πρόβλεψη; Βρείτε έναν τρόπο (ποσοτικό ή ποιοτικό) για να το αιτιολογήσετε.

2 Εφαρμογή σε πρόβλημα ταξινόμησης σκηνής

Ένα από τα πιο δημοφιλή προβλήματα στο χώρο της σημασιολογικής ανάλυσης πολυμεσικού υλικού αποτελεί η ταξινόμηση εικόνων με βάση το περιεχόμενο. Πιο συγκεκριμένα, ο σκοπός είναι η ταξινόμηση με βάση την σκηνή ή τα αντικείμενα που περιέχουν οι εικόνες. Σε τέτοια προβλήματα, η διαδικασία που ακολουθείται για την επίλυσή τους είναι η εξής:

- Επιλογή κατάλληλων οπτικών χαρακτηριστικών χαμηλού επιπέδου (π.χ. χρώμα, υφή) για την επίλυση του προβλήματος.
- Επιλογή ενός κατάλληλου συνόλου εικόνων για τη χρησιμοποίησή τους σαν σύνολο εκμάθησης.
- Εξαγωγή των επιλεχθέντων οπτικών χαρακτηριστικών από ολόκληρη την εικόνα, ή από μέρη αυτής, για όλες τις εικόνες του συνόλου εκμάθησης.
- Εκπαίδευση κατάλληλου ταξινομητή με βάση τα οπτικά χαρακτηριστικά που επιλέχθηκαν.
- Επιλογή εικόνων για τη χρησιμοποίησή τους σαν σύνολο ελέγχου.
- Έλεγχος της επίδοσης του ταξινομητή με βάση κατάλληλα μέτρα.
- Εκ νέου εκπαίδευση ή/και επιλογή διαφορετικών οπτικών χαρακτηριστικών στην περίπτωση που τα αποτελέσματα είναι κατώτερα από τα επιθυμητά.

Στο μέρος αυτό της άσκησης, να κατασκευάσετε ένα νευρωνικό δίκτυο που θα λειτουργεί ως ταξινομητής εικόνων με βάση τη θεματική κατηγορία στην οποία ανήκουν. Η 1η θεματική κατηγορία είναι “Αρχαία Μνημεία” και η 2η “Φυσικό Περιβάλλον”. Το σύνολο εκμάθησης περιέχει 200 εικόνες από την κάθε θεματική κατηγορία και το σύνολο ελέγχου περιέχει 78 εικόνες από την 1η και 55 εικόνες από τη δεύτερη. Πιο συγκεκριμένα, να ακολουθήσετε τα παρακάτω βήματα:

1. Να εξάγετε μια περιγραφή χρώματος από το σύνολο των εικόνων με τη χρήση της συνάρτησης `getImageDescription` που παρέχεται. Η συνάρτηση αυτή διαβάζει μια εικόνα και την αποθηκεύει στο χώρο εργασίας της MATLAB. Στην συνέχεια γίνεται μετατροπή της εικόνας από τον RGB στον HSV χρωματικό χώρο, και για κάθε συνιστώσα του χώρου εξάγεται ένα ιστόγραμμα. Το ιστόγραμμα αυτό δείχνει την κατανομή των τιμών των pixels της εικόνας. Στη συνέχεια τα 3 ιστογράμματα “ενώνονται” και προκύπτει μοναδική περιγραφή για κάθε εικόνα.

Η συνάρτηση `getImageDescription` δέχεται σαν όρισμα το όνομα μιας εικόνας και των αριθμό των κορυφών για κάθε συνιστώσα του χώρου HSV.

2. Να επιλέξετε κατάλληλη αρχιτεκτονική και παραμέτρους και να εκπαιδεύσετε ένα νευρωνικό δίκτυο που να επιλύει το πρόβλημα, με βάση την εμπειρία που αποκτήσατε στο προηγούμενο μέρος της άσκησης. Να θέσετε ένα κατώφλι πάνω από το οποίο ή έξοδος του δικτύου θα αντιστοιχεί στη μία κατηγορία και κάτω από το οποίο στην άλλη.
3. Να μελετήσετε το πώς η αρχιτεκτονική του δικτύου, αλλά και η λεπτομέρεια της περιγραφής (δηλαδή ο αριθμός των κορυφών του ιστογράμματος) επηρεάζει την επίδοση του ταξινομητή. Για την επίδοση του ταξινομητή χρησιμοποιήστε κάποιο από τα ακόλουθα μέτρα:

$$Precision = \frac{tp}{tp + fp} \quad (1)$$

$$Accuracy = \frac{tp + tn}{tp + fp + fn + tn} \quad (2)$$

Όπου όταν αναφερόμαστε σε μια θεματική κατηγορία, tp είναι οι εικόνες που ανιχνεύθηκαν ότι ανήκουν στην κατηγορία αυτή, ενώ πράγματι ανήκουν (true positives), tn είναι οι εικόνες που ανιχνεύθηκαν ότι δεν ανήκουν στην κατηγορία αυτή, ενώ πράγματι δεν ανήκουν (true negatives), fp είναι οι εικόνες που ανιχνεύθηκαν ότι ανήκουν στην κατηγορία αυτή, ενώ δεν ανήκουν (false positives) και fn είναι οι εικόνες που ανιχνεύθηκαν ότι δεν ανήκουν στην κατηγορία αυτή, ενώ ανήκουν (false negatives).

4. Ανάμεσα στις εικόνες που χρησιμοποιήσατε, υπάρχουν και κάποιες οι οποίες δεν είναι δυνατόν να ταξινομηθούν σωστά. Προσπαθήστε να το εξηγήσετε παρατηρώντας τα οπτικά τους χαρακτηριστικά.

A Υποδείξεις

A.1 Γενικά, για όλη την άσκηση

- α. Να γράψετε ένα ξεχωριστό script για κάθε ερώτημα, για να διευκολύνετε την επίδειξη και τη διόρθωση της άσκησης. Τα scripts αυτά θα τα παραδώσετε σε ηλεκτρονική μορφή. Να φροντίσετε να είναι ευανάγνωστα και με τα απαραίτητα σχόλια.
- β. Για να λάβετε πιο αξιόπιστα αποτελέσματα, μπορείτε κάθε φορά να επαναλαμβάνετε κάθε πείραμα 5 φορές και να παίρνετε τον μέσο όρο των αποτελεσμάτων.
- γ. Μπορείτε να πειραματιστείτε (προαιρετικά) και με τις ακόλουθες συναρτήσεις εκπαίδευσης της MATLAB: `traingdx`, `trainlm`, `trainrp`, `trainbfg` και `traincgb`. Να σχολιάσετε συνοπτικά τις παρατηρήσεις σας.

A.2 Ειδικά, ανά ερώτημα

- 1.1.1 Τα σύνολα εκμάθησης (training) και εκπαίδευσης (test) δίνονται με τη μορφή των δύο m-files `Traindata.m` και `Testdata.m`. Εκτελώντας τα m-files αυτά, φορτώνονται στο workspace της MATLAB οι πίνακες `Traindata` και `Testdata`. Στους πίνακες αυτούς, κάθε γραμμή αντιπροσωπεύει ένα παράδειγμα και κάθε στήλη ένα χαρακτηριστικό. Η τελευταία στήλη αντιπροσωπεύει την επιθυμητή απόκριση κάθε παραδείγματος.
- 1.1.3 Για την προεπεξεργασία των δεδομένων, να χρησιμοποιήσετε τις συναρτήσεις `premmmx`, `prestd`, `prepca`. Μετά την εκπαίδευση του δικτύου, να φέρετε τα δεδομένα στην κατάλληλη μορφή με τη χρήση των `postmmmx`, `poststd`. Είναι προφανές ότι πρέπει να κάνετε προεπεξεργασία και στα δεδομένα που θα χρησιμοποιήσετε για να ελέγξετε το δίκτυο. Χρησιμοποιήστε κατάλληλα τις συναρτήσεις `trastd`, `trapca` και `trammmx`. Σε καινούριες

εκδόσεις της MATLAB οι συναρτήσεις είναι obsolete, αλλά μπορείτε να τις χρησιμοποιήσετε κανονικά.

- 1.2.1 Για να κατασκευάσετε ένα feed-forward νευρωνικό δίκτυο, να χρησιμοποιήσετε την έτοιμη συνάρτηση της MATLAB `newff` για τη δομή του δικτύου και την `traingd` σαν συνάρτηση εκπαίδευσης. Απαιτούνται προφανώς τόσες είσοδοι στο δίκτυο, όση είναι και η διάσταση του σύνολου εκπαίδευσης **μετά** την PCA, καθώς και μία έξοδος η οποία θα δίνει την εκτίμηση της τιμής του σπιτιού. Στην αναφορά, να παραθέσετε γραφική παράσταση που να δείχνει την μεταβολή του μέσου τετραγωνικού σφάλματος σε σχέση με τον αριθμό των νευρώνων.
 - 1.2.2 Για προσαρμοζόμενο ρυθμό μάθησης, να χρησιμοποιήσετε για την εκπαίδευση του δικτύου τη συνάρτηση `traingda`. Στην αναφορά, να παραθέσετε τις καμπύλες μάθησης για κάθε περίπτωση ρυθμού μάθησης που θα σχολιάσετε ποιοτικά, καθώς και το μέσο τετραγωνικό σφάλμα κάθε δικτύου.
 - 1.2.3 Για την εκπαίδευση του δικτύου με τη χρήση όρου ορμής, να χρησιμοποιήσετε σαν συνάρτηση εκπαίδευσης την `traingdm`. Να παραθέσετε και να σχολιάσετε κατάλληλη καμπύλη μάθησης. Πώς το ερμηνεύεται αυτό θεωρητικά;
 - 1.2.4 Για την μέθοδο “Early Stopping of Training” να ορίσετε κατάλληλα μια δομή που να περιλαμβάνει το validation training set και το validation test set που επιλέξατε και να την δώσετε σαν όρισμα στην `train`. Μια “καλή” αναλογία των παραπάνω συνόλων προς τα σύνολα εκπαίδευσης και εκμάθησης είναι 30%.
- 2.1 Τη συνάρτηση `getImageDescription`, καθώς και τις εικόνες που αποτελούν τα σύνολα εκμάθησης και ελέγχου (περίπου 16 MB) μπορείτε να τα βρείτε στο site του μαθήματος, μαζί και με το υπόλοιπο υλικό της άσκησης.
 - 2.3 Να πειραματιστείτε για αρκετές τιμές των κορυφών του ιστογράμματος (π.χ. 2,4,8,16,...,256) και να παρουσιάσετε τα αποτελέσματα με τη μορφή κατάλληλης γραφικής παράστασης.

B Επικοινωνία

Το site του μαθήματος είναι: www.image.ece.ntua.gr/courses/nn. Για ο,τιδήποτε χρειαστείτε σχετικά με την 1η εργαστηριακή άσκηση, μπορείτε να απευθύνεστε στους:

- Νίκος Σίμου, nsimou@image.ece.ntua.gr, τηλ. 210-7723039
- Βαγγέλης Σπύρου, espyrou@image.ece.ntua.gr, τηλ. 210-7722491