

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ
Χειμερινό Εξάμηνο 2016-17

1η Εργαστηριακή Άσκηση: Οπτική Αναγνώριση Χαρακτήρων

ΠΕΡΙΓΡΑΦΗ

Σκοπός είναι η υλοποίηση ενός συστήματος οπτικής αναγνώρισης ψηφίων. Τα δεδομένα προέρχονται από την US Postal Service (γγραμμένα στο χέρι σε ταχυδρομικούς φακέλους και σκαναρισμένα) και περιέχουν τα ψηφία από το 0 έως το 9 και διακρίνονται σε train και test.

Τα δεδομένα κάθε αρχείου αναπαριστούν τα περιεχόμενα ενός πίνακα (οι τιμές των στοιχείων του πίνακα διαχωρίζονται με κενό). Κάθε γραμμή αφορά ένα ψηφίο (δείγμα). Οι στήλες αντιστοιχούν στα χαρακτηριστικά (features) που περιγράφουν τα ψηφία. Για παράδειγμα, η τιμή του (i,j) στοιχείου αφορά το j -th χαρακτηριστικό του i -th ψηφίου. Κάθε ψηφίο περιγράφεται από 257 τιμές, εκ των οποίων η πρώτη αντιστοιχεί στο ίδιο το ψηφίο (αν είναι το 0, το 1 κτλ.) και οι υπόλοιπες 256 είναι τα χαρακτηριστικά (features) που το περιγράφουν (grayscale values). Ας φανταστούμε το κάθε ψηφίο να απεικονίζεται σε έναν 16×16 πίνακα αποτελούμενο από 256 κουτάκια ("pixels"). Για να εμφανίζεται το κάθε ψηφίο στην οθόνη "φωτίζεται" ένα σύνολο από τέτοια κουτάκια, με τέτοιο τρόπο ώστε η συνολική εικόνα που βλέπουμε να απεικονίζει το θεωρούμενο ψηφίο. Επειδή τα ψηφία εμφανίζονται σε grayscale, κάθε μία από τις 256 τιμές αντιστοιχεί σε μία απόχρωση μαύρου για το αντίστοιχο "pixel". Στόχος είναι η δημιουργία και αποτίμηση (evaluation) ταξινομητών οι οποίοι θα ταξινομούν κάθε ένα από τα ψηφία που περιλαμβάνονται στα test δεδομένα σε μία από τις δέκα κατηγορίες (από το 0 έως το 9).

ΕΚΤΕΛΕΣΗ

Χρησιμοποιώντας τα δεδομένα που κατεβάσατε ήδη (κατά την προπαρασκευή, από το link https://www.dropbox.com/s/gpnp72vw22q3j3a/pr_lab1_2016-17_pswd.zip?dl=0) καθώς και τον κώδικα που αναπτύξατε κατά την προπαρασκευή, εκτελέστε τα παρακάτω βήματα χρησιμοποιώντας Matlab:

Σημείωση: Τα βήματα 1–9 αποτελούν μέρος της προπαρασκευής η οποία έχει προηγηθεί και επαναλαμβάνονται για σκοπούς πληρότητας της άσκησης.

Βήμα 1

Σχεδιάστε το υπ' αριθμόν 131 ψηφίο (βρίσκεται στη θέση 131) των train δεδομένων.

Βήμα 2

Υπολογίστε τη μέση τιμή των χαρακτηριστικών του pixel (10, 10) για το ψηφίο «0» με βάση τα train δεδομένα.

Βήμα 3

Υπολογίστε τη διασπορά των χαρακτηριστικών του pixel (10, 10) για το ψηφίο «0» με βάση τα train δεδομένα.

Βήμα 4

Υπολογίστε τη μέση τιμή και διασπορά των χαρακτηριστικών κάθε pixel για το ψηφίο «0» με βάση τα train δεδομένα.

Βήμα 5

Σχεδιάστε το ψηφίο «0» χρησιμοποιώντας τις τιμές της μέσης τιμής που υπολογίσατε στο Βήμα 4.

Βήμα 6

Σχεδιάστε το ψηφίο «0» χρησιμοποιώντας τις τιμές της διασποράς που υπολογίσατε στο Βήμα 4. Συγκρίνετε το αποτέλεσμα με το αποτέλεσμα του Βήματος 5 και εξηγήστε τυχόν διαφορές.

Βήμα 7

(α) Υπολογίστε τη μέση τιμή και διασπορά των χαρακτηριστικών για όλα τα ψηφία (0-9) με βάση τα train δεδομένα.

(β) Σχεδιάστε όλα τα ψηφία χρησιμοποιώντας τις τιμές της μέσης τιμής που υπολογίσατε στο Βήμα 7(α).

Βήμα 8

Ταξινομήστε το υπ' αριθμόν 101 ψηφίο των test δεδομένων ως προς τις 10 κατηγορίες (κάθε ένα από τα 10 ψηφία, 0-9, αντιπροσωπεύει μία κατηγορία) βάσει της Ευκλείδειας απόστασης (χρησιμοποιείτε τις τιμές που υπολογίσατε στο Βήμα 7(α)).

Βήμα 9

(α) Ταξινομήστε όλα τα ψηφία των test δεδομένων ως προς τις 10 κατηγορίες με βάση την Ευκλείδεια απόσταση.

(β) Υπολογίστε το ποσοστό επιτυχίας για το Βήμα 9(α).

Τα παρακάτω βήματα δεν αποτελούν μέρος της προπαρασκευής.

Βήμα 10

Υπολογίστε τις a-priori πιθανότητες για κάθε κατηγορία.

Βήμα 11

(α) Ταξινομήστε όλα τα ψηφία των test δεδομένων ως προς τις 10 κατηγορίες χρησιμοποιώντας τις τιμές της μέσης τιμής και διασποράς που υπολογίσατε στο Βήμα 7(α), υλοποιώντας έναν Bayesian ταξινομητή.

Υπόδειξη: μην χρησιμοποιήσετε έτοιμες υλοποιήσεις.

(β) Υπολογίστε το ποσοστό επιτυχίας για το Βήμα 11(α).

Βήμα 12

Επαναλάβετε το Βήμα 11(α), (β) υποθέτοντας ότι η διασπορά για όλα τα χαρακτηριστικά, για όλες τις κατηγορίες ισούται με 1.

Βήμα 13

(α) Ταξινομήστε τα 100 πρώτα ψηφία των test δεδομένων χρησιμοποιώντας τα 1000 πρώτα ψηφία των train δεδομένων. Ως αλγόριθμο ταξινόμησης εφαρμόστε τον αλγόριθμο του Κοντινότερου Γείτονα-1 (Nearest Neighbor-1 (NNR-1)).

Υπόδειξη: μην χρησιμοποιήσετε έτοιμες υλοποιήσεις.

(β) Υπολογίστε το ποσοστό επιτυχίας για το Βήμα 13(α).

Βήμα 14

(α) Ταξινομήστε όλα τα ψηφία των test δεδομένων χρησιμοποιώντας όλα τα ψηφία των train δεδομένων. Ως αλγόριθμο ταξινόμησης εφαρμόστε τον αλγόριθμο του Κοντινότερου Γείτονα-1 (Nearest Neighbor-1 (NNR-1)).

(β) Υπολογίστε το ποσοστό επιτυχίας για το Βήμα 14(α).

(γ) Επαναλάβετε τα Βήματα 14(α), (β) για NNR-k (αντί του NNR-1). Δοκιμάστε διάφορες τιμές του k.

Υπόδειξη: χρησιμοποιείτε περιττό αριθμό για τις τιμές του k.

(δ) Προτείνετε μία βελτιωμένη παραλλαγή του NNR-k και δοκιμάστε την.

Βήμα 15

(α) Αφήνεται ως ελεύθερη επιλογή η εφαρμογή άλλων ταξινομητών (*) για την ταξινόμηση των ψηφίων. Επιλέξτε τουλάχιστον δύο ταξινομητές συμπεριλαμβάνοντας σε αυτούς τα Support Vector Machines (SVM). Κατά την εφαρμογή των SVM χρησιμοποιήστε δύο τύπους συναρτήσεων για τα kernels: (i) γραμμικούς, και (ii) πολυωνυμικούς. Υπόδειξη: μπορείτε να χρησιμοποιήσετε έτοιμες υλοποιήσεις, π.χ., Matlab `svmtrain/svmclassify` (**) για SVM. (Σε περίπτωση που εφαρμόσετε έτοιμες υλοποιήσεις αναφέρατε την πηγή σας και τις συγκεκριμένες παραμέτρους που χρησιμοποιούνται.)

(*) Μπορείτε να χρησιμοποιήσετε και εργαλεία εκτός Matlab. Για παράδειγμα, αναφορικά με SVM υπάρχουν διάφορα δημοφιλή εργαλεία όπως SVM Light (<http://svmlight.joachims.org/>), LIBSVM (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>), OSU SVM (<http://svm.sourceforge.net/license.shtml>).

(**) Οι συγκεκριμένες συναρτήσεις του Matlab αφορούν SVM για προβλήματα 2 κατηγοριών. Μια απλή τεχνική που χρησιμοποιείται σε προβλήματα με περισσότερες κατηγορίες δίνεται μέσω του εξής παραδείγματος. Εστω ένα πρόβλημα κατηγοριοποίησης με 3 κατηγορίες. Αναπτύσσονται 3 ταξινομητές: (i) κατηγορία 1 vs υπόλοιπες (2 και 3), (ii) κατηγορία 2 vs υπόλοιπες (1 και 3), και (iii) κατηγορία 3 vs υπόλοιπες (1 και 2). Για την κατηγοριοποίηση ενός άγνωστου ψηφίου χρησιμοποιούνται και οι 3 ταξινομητές. Το άγνωστο ψηφίο κατηγοριοποιείται σε μία κατηγορία λαμβάνοντας υπόψη τις κατηγοριοποιήσεις των 3 ταξινομητών (πώς;).

(β) Υπολογίστε το ποσοστό επιτυχίας για τους διάφορους ταξινομητές και σχολιάστε τα αποτελέσματα. Στο σχολιασμό συμπεριλάβετε και τους ταξινομητές των προηγούμενων βημάτων.

Βήμα 16

Η βασική ιδέα του βήματος αυτού είναι ο συνδυασμός κάποιων ταξινομητών με αρκετά καλή επίδοση με στόχο να επιτευχθεί επίδοση υψηλότερη των επιμέρους επιδόσεων. Είναι σημαντικό οι ταξινομητές που θα συνδυαστούν να χαρακτηρίζονται από διαφορετικό τύπο λαθών, π.χ., ο ένας ταξινομητής να τείνει να ταξινομεί λάθος το ψηφίο 3, ενώ ένας άλλος να τείνει να ταξινομεί λάθος το ψηφίο 7. Επιλέξτε κάποιους από τους ταξινομητές που χρησιμοποιήσατε στα προηγούμενα βήματα.

(α) Ένα απλός συνδυασμός είναι να ληφθεί υπόψη η πλειοψηφία των επί μέρους κατηγοριοποιήσεων. Δηλαδή, ένα άγνωστο ψηφίο ταξινομείται στην κατηγορία εκείνη στην οποία γίνεται κατηγοριοποίηση από τους περισσότερους ταξινομητές. Υπόδειξη: χρησιμοποιήστε περιττό πλήθος ταξινομητών. Για παράδειγμα, υποθέστε ότι έχουμε 2 κατηγορίες και 3 ταξινομητές. Έστω ότι ένα άγνωστο ψηφίο κατηγοριοποιείται στην κατηγορία 1, 2, και 1 από τον 1ο, 2ο, και 3ο ταξινομητή, αντίστοιχα. Λαμβάνοντας υπόψη την πλειοψηφία, το άγνωστο ψηφίο κατηγοριοποιείται τελικώς στην κατηγορία 1. Τί θα κάνατε σε περίπτωση που δεν είναι δυνατό να βρεθεί πλειοψηφική κατηγοριοποίηση;

(β) Για κάθε άγνωστο ψηφίο και ως προς κάθε υποψήφια κατηγορία, ένας ταξινομητής υπολογίζει τα αντίστοιχα confidence scores (*) ή εκτιμήσεις των posterior πιθανοτήτων. Σκοπός είναι ο συνδυασμός αυτών των scores για τον υπολογισμό ενός νέου, συνολικού score για κάθε άγνωστο ψηφίο ως προς κάθε κατηγορία, βάσει του οποίου γίνεται η τελική ταξινόμηση. Για παράδειγμα, υποθέστε ότι έχουμε 2 κατηγορίες και 3 ταξινομητές. Έστω ότι τα scores του 1ου ταξινομητή είναι [0.3 0.5], του 2ου ταξινομητή [0.1 0.9] και του 3ου ταξινομητή [0.4 0.6] ως προς την κατηγορία 1 και 2, αντίστοιχα. Μπορούμε να καταλήξουμε σε ένα συνολικό score για κάθε κατηγορία εφαρμόζοντας μία συνάρτηση των επιμέρους scores. Χρησιμοποιώντας για παράδειγμα τον πολλαπλασιασμό, καταλήγουμε στο [0.012 0.270] το οποίο υποδεικνύει την κατηγοριοποίηση στην κατηγορία 2. Δοκιμάστε διάφορες συναρτήσεις όπως min, max, αριθμητικό μέσο, και πολλαπλασιασμό. Χωρίστε τα train δεδομένα σε δυο υποσύνολα (π.χ., να περιέχουν το 80% και 20% των δεδομένων, αντίστοιχα) προσπαθώντας δεδομένα για κάθε κατηγορία να περιέχονται και στα δύο υποσύνολα. Το πρώτο υποσύνολο χρησιμοποιείται για την εκπαίδευση των επιμέρους ταξινομητών και τον υπολογισμό των αντίστοιχων confidence scores/εκτιμήσεις posterior πιθανοτήτων. Αποτιμήστε τις προαναφερθείσες συναρτήσεις (min, max, κτλ) χρησιμοποιώντας το δεύτερο υποσύνολο. Για τη συνάρτηση με την υψηλότερη επίδοση χρησιμοποιήστε τα test δεδομένα για τελική αποτίμηση.

(*) Σε περίπτωση που η υλοποίηση που χρησιμοποιήσατε δεν δίνει στην έξοδο της κάποιο confidence score, μπορείτε να ορίσετε εσείς το εν λόγω score δίνοντας τις κατάλληλες εξηγήσεις.

(γ) Θεωρήστε τα confidence scores ή τις εκτιμήσεις των posterior πιθανοτήτων που υπολογίζουν οι επιμέρους ταξινομητές ως χαρακτηριστικά. Χρησιμοποιήστε αυτά τα νέα χαρακτηριστικά σε έναν ταξινομητή (υπόδειξη: μπορείτε να επιλέξετε έναν ταξινομητή που προβαίνει σε γραμμικό συνδυασμό χαρακτηριστικών).

Χρησιμοποιήστε τα δυο υποσύνολα των train δεδομένων που φτιάξατε στο προηγούμενο βήμα. Το πρώτο υποσύνολο χρησιμοποιείται για την εκπαίδευση των επιμέρους ταξινομητών. Το δεύτερο υποσύνολο χρησιμοποιείται για την εκπαίδευση του ταξινομητή ο οποίος δέχεται ως χαρακτηριστικά τα προαναφερθέντα confidence scores/εκτιμήσεις posterior πιθανοτήτων. Η επίδοση του τελευταίου ταξινομητή υπολογίζεται ως προς τα test δεδομένα.

(δ) Για κάθε ένα από τα βήματα (α) – (γ) αναφέρατε το ποσοστό επιτυχίας της ταξινόμησης. Επίσης, συγκρίνετε και σχολιάστε τους διάφορους συνδυασμούς. Τέλος, προτείνετε επιπλέον συνδυασμούς.

Προχωρήστε σε κατ' οίκον ολοκλήρωση των βημάτων εκείνων που δεν προλάβετε κατά τη διεξαγωγή του εργαστηρίου.

ΠΑΡΑΔΟΤΕΑ

Αφορά τα εκτός προπαρασκευής βήματα μόνο.

(1) Σύντομη αναφορά (σε pdf) που θα περιγράφει τη διαδικασία που ακολουθήθηκε σε κάθε βήμα, καθώς και τα σχετικά αποτελέσματα.

(2) Κώδικας Matlab (συνοδευόμενος από σύντομα σχόλια).

Συγκεντρώστε τα (1) και (2) σε ένα .zip αρχείο το οποίο πρέπει να αποσταλεί μέσω του mycourses.ntua.gr εντός της καθορισμένης προθεσμίας.