

ΕΠ08 Αναγνώριση Προτύπων – Μηχανική Μάθηση

2^η Εργασία

Τύπος εργασίας: **Ατομική**

Ημερομηνία παράδοσης: **Κυριακή 21/05/2023, 23:55 (Δεν θα δοθεί παράταση)**

Τρόπος παράδοσης: **Αποκλειστικά μέσω του eclass**

Σύνολο βαθμών: **100** (15% του τελικού βαθμού του μαθήματος)

Η εργασία είναι **ατομική** και αποτελείται από 2 ερωτήματα. Συνιστάται ιδιαίτερα, να αφιερώσετε χρόνο ώστε να κατανοήσετε τη λογική πίσω από τα ερωτήματα της εργασίας και να αποφύγετε την αναζήτηση έτοιμων λύσεων στο διαδίκτυο. Αν ωστόσο συμβουλευτείτε ή/και χρησιμοποιήσετε οποιοδήποτε υλικό ή/και κώδικα που είναι διαθέσιμος στο διαδίκτυο, πρέπει να αναφέρεται σωστά τη πηγή ή/και το σύνδεσμο στην ιστοσελίδα που αντλήσατε πληροφορίες. Σε κάθε περίπτωση, η αντιγραφή τμήματος ή του συνόλου της εργασίας δεν είναι αποδεκτή και στη περίπτωση που διαπιστωθεί αντιγραφή θα μηδενιστούν στο μάθημα όλα τα εμπλεκόμενα μέρη.

Θα πρέπει να υποβάλετε **ένα μόνο αρχείο Notebook IPython (Jupyter notebook) μέσω του εργαλείου εργασίες του eclass**, ακολουθώντας την εξής σύμβαση ονομασίας για το αρχείο σας: **Επώνυμο_ΑριθμόςΜητρώου.ipynb**

Μπορείτε να χρησιμοποιήσετε κελιά επικεφαλίδας για να οργανώσετε περαιτέρω το έγγραφό σας. **Σημαντικό:** Το IPython notebook που θα παραδώσετε θα πρέπει βεβαιωθείτε ότι ανοίγει και να εκτελείται στο google colab.

[Ερώτημα 1: Αναγνώριση Προσώπων (Face recognition)] (70 βαθμοί)

Σε αυτό το ερώτημα θα εφαρμόσετε τη μέθοδο Eigenfaces (δηλαδή συνδυασμό PCA για εξαγωγή χαρακτηριστικών και ταξινομητή πλησιέστερου γείτονα για την αναγνώριση προσώπων). Θα χρησιμοποιήσετε εικόνες προσώπων από τη βάση δεδομένων προσώπων Yale B στην οποία υπάρχουν 10 πρόσωπα που φωτογραφήθηκαν κάτω από 64 διαφορετικές συνθήκες φωτισμού. Χρησιμοποιώντας την υλοποίησή σας, θα αξιολογήσετε την ικανότητα του αλγορίθμου Eigenfaces να χειρίζεται συνθήκες φωτισμού των εικόνων ελέγχου (test set) οι οποίες διαφέρουν από αυτές στις εικόνες εκπαίδευσης (training set).

Τα δεδομένα είναι διαθέσιμα στο αρχείο faces.zip στο κατάλογο Έγγραφα στο eclass.

Η μέθοδος Eigenfaces για την αναγνώριση προσώπων περιλαμβάνει 3 βασικά βήματα:

Βήμα 1: Κάθε εικόνα διάστασης 50 x 50 pixels του συνόλου εκπαίδευσης μετατρέπεται σε διάνυσμα διάστασης 2500 στοιχείων και αποθηκεύεται ως στήλη στον πίνακα δεδομένων εκπαίδευσης X . Στη συνέχεια εφαρμόζουμε principal component analysis (PCA) στον πίνακα δεδομένων εκπαίδευσης και εξάγουμε τις d κύριες συνιστώσες (principal components). Τα d ιδιοδιανύσματα (eigenvectors) όταν μετατραπούν και απεικονιστούν ως εικόνες ονομάζονται Eigenfaces.

Βήμα 2: Προβάλλουμε τις εικόνες των συνόλων εκπαίδευσης και ελέγχου στο χώρο d διαστάσεων και με αυτόν το τρόπο εξάγουμε χαρακτηριστικά χαμηλής διάστασης (d -dimensional features). Ο χώρος χαμηλής διάστασης d ονομάζεται ιδιοχώρος (eigenspace).

Βήμα 3: Η αναγνώριση των προσώπων γίνεται στον eigenspace χρησιμοποιώντας ταξινομητή (ενός) πλησιέστερου γείτονα με Ευκλείδεια απόσταση ως μετρική.

Από το σύνολο δεδομένων προσώπων Yale B θα χρησιμοποιήσετε τα παρακάτω υποσύνολα:

Set_1: person*_01.png έως person*_07.png (δηλαδή τις 7 πρώτες εικόνες κάθε προσώπου)

Set_2: person*_08.png έως person*_19.png

Set_3: person*_20.png έως person*_31.png

Set_4: person*_32.png έως person*_45.png

Set_5: person*_46.png έως person*_64.png

Ζητούμενα:

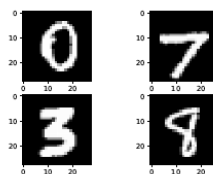
- I. Να γράψετε μία συνάρτηση loadImages(path, set_number) η οποία παίρνει ως είσοδο το path στο οποίο βρίσκεται ο φάκελος των εικόνων π.χ. loadImages("C:/images", "Set_1"), διαβάζει τις εικόνες και επιστέφει έναν πίνακα δεδομένων ανάλογα με το set_number, όπου κάθε εικόνα αναπαρίσταται ως διάνυσμα στήλη. Η συνάρτηση επιστέφει επίσης τις κατηγορίες (labels) στις οποίες ανήκουν οι διαφορετικές εικόνες κωδικοποιημένες με ακραίους (π.χ. 0 για φωτογραφίες που ανήκουν στο person_0, 1 για τις φωτογραφίες που ανήκουν στο person_1 κτλ).
- II. Να εκπαιδεύσετε την μέθοδο Eigenfaces με $d = 9$ και $d = 30$ χρησιμοποιώντας όλες τις εικόνες στο Set_1 (70 εικόνες) και να αναγνωρίσετε τα πρόσωπα στα Set_1 έως Set_5. Για κάθε Set και κάθε τιμή της διάστασης d να αναφέρετε την ακρίβεια ταξινόμησης. Για το Set_1 αναμένουμε 100% ακρίβεια ταξινόμησης καθώς χρησιμοποιήθηκε για την εκπαίδευση της μεθόδου Eigenfaces. Σχολιάστε την δυνατότητα γενίκευσης της μεθόδου στα διαφορετικά Sets.

- III. Να απεικονίσετε (σε μορφή εικόνας) τα 9 κύρια ιδιοδιανύσματα (9 top eigenvectors) που προέκυψαν αφού εκπαιδεύσατε την μέθοδο Eigenfaces στο Set_1. Τι παρατηρείτε; Τι θα μπορούσαμε να πούμε ότι εκφράζουν τα διαφορετικά ιδιοδιανύσματα;
- IV. Να χρησιμοποιήσετε $d = 9$ και $d = 30$ Eigenfaces που βρήκατε από το Set_1, για να ανακατασκευάσετε μια τυχαία εικόνα από κάθε ένα από τα 5 Sets. Να απεικονίσετε τόσο τις αρχικές εικόνες όσο και τις ανακατασκευασμένες της για διαφορετικές τιμές του d . Να σχολιάσετε την ποιότητα ανακατασκευής κάθε εικόνας.
- V. Να απεικονίσετε τα 9 κύρια singular vectors που προκύπτουν αφού εφαρμόσετε SVD στον πίνακα δεδομένων του Set_1. Διαφέρουν τα singular vectors από τα αντίστοιχα ιδιοδιανύσματα; Αν ναι, γιατί;

Σημείωση: Μπορείτε να χρησιμοποιήσετε είτε έτοιμες υλοποιήσεις της PCA είτε να την υλοποιήσετε χρησιμοποιώντας ιδιοανάλυση (συναρτήσεις τύπου eig) στον πίνακα συνδιακύμανσης. Προτείνεται να προ-επεξεργαστείτε κάθε εικόνα αφαιρώντας τη μέση τιμή της και διαιρώντας με την τυπική απόκλιση των τιμών της.

[Ερώτημα 2: Ταξινόμηση εικόνων χρησιμοποιώντας SVMs] (30 βαθμοί)

Σε αυτό το ερώτημα, καλείστε να διερευνήσετε την επίδοση των support vector machines στο πρόβλημα της αναγνώρισης χειρόγραφων ψηφίων. Για το σκοπό αυτό θα χρησιμοποιήσετε το σύνολο δεδομένων MNIST και υλοποιήσεις αλγορίθμων της βιβλιοθήκης scikit-learn. Το σύνολο δεδομένων MNIST αποτελείται από 70000 εικόνες χειρόγραφων ψηφίων και, τυπικά, χωρίζεται σε τρία υποσύνολα: training set (50000 εικόνες), validation set (10000 εικόνες), test set (10000 εικόνες). Κάθε εικόνα έχει διάσταση 28×28 pixels και απεικονίζει ένα χειρόγραφο ψηφίο. Παραδείγματα τέτοιων εικόνων απεικονίζονται στο παρακάτω σχήμα:



- Ζητείται να φορτώσετε τα δεδομένα του συνόλου MNIST και να μετατρέψετε κάθε εικόνα σε μορφή διανύσματος διάστασης $28 \times 28 = 784$. Στη συνέχεια κανονικοποιήστε (normalize) τα δεδομένα στο διάστημα $[0,1]$.
- Στα SVMs υπάρχουν διάφορες επιλογές που μπορεί να επηρεάσουν την απόδοση τους σε προβλήματα ταξινόμησης. Παραδείγματα τέτοιων επιλογών αποτελούν ο τύπος του πυρήνα (kernel) και οι τιμές των διάφορων (υπερ)παραμέτρων. Ζητείται να εξετάσετε την επίδοση των SVMs για γραμμικό (linear SVMs) και RBF πυρήνα και διαφορετικές τιμές παραμέτρων ώστε να καθορίσετε το συνδυασμό παραμέτρων/πυρήνων που οδηγούν στη μεγαλύτερη ακρίβεια ταξινόμησης. Για

αυτό το πείραμα να χρησιμοποιήσετε 60000 εικόνες για εκπαίδευση (training) και 10000 παραδείγματα για δοκιμές (test). Να αναφέρετε τις τιμές των παραμέτρων, δηλαδή τύπο πυρήνα, τιμές των C και γ που οδηγούν στις καλύτερες επιδόσεις τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο δοκιμής (test set).

- Στη συνέχεια, να εφαρμόσετε PCA στα δεδομένα επιλέγοντας 3 διαφορετικές τιμές για τη διατηρούμενη διακύμανση και για κάθε τιμή διακύμανσης εκτελέστε ξανά τη μέθοδο SVM χρησιμοποιώντας τις παραμέτρους που οδήγησαν στην καλύτερη επίδοση στο παραπάνω ερώτημα. Για κάθε εκτέλεση, αναφέρετε τον αριθμό των συνιστωσών (components) που διατηρούνται καθώς και την ακρίβειας ταξινόμησης. Επίσης, καταγράψτε τους χρόνους εκτέλεσης κάθε πειράματος και εξαγάγετε συμπεράσματα σχετικά με μια πιθανή αντιστάθμιση (trade-off) μεταξύ ακρίβειας ταξινόμησης, μείωσης διαστάσεων και χρόνου εκτέλεσης του αλγορίθμου.