

**Τμήμα Μηχανικών Η/Υ & Πληροφορικής,
Πανεπιστήμιο Ιωαννίνων**

Μεταπτυχιακό μάθημα: «Μηχανική Μάθηση»

(Ημερομηνία παράδοσης: έως 10/6/2019)

Θέμα: Μείωση διάστασης και εφαρμογή στην αναγνώριση προσώπων

Στην ιστοσελίδα:

<https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

μπορείτε να βρείτε ένα σύνολο αποτελούμενο από 400 εικόνες 40 προσώπων (10 διαφορετικές εικόνες για κάθε πρόσωπο) με διαφορετικές όψεις, φωτεινότητες, κλπ. Διαβάστε σχετικές πληροφορίες στην σελίδα. Θα πρέπει να κάνετε τα εξής:

(α) Ταξινόμηση Προσώπων: Επιλέξτε για κάθε μία κατηγορία-πρόσωπο τις μισές ($40 \times 5 = 200$) εικόνες για εκπαίδευση και τις υπόλοιπες (200) για έλεγχο. Στο σύνολο εκπαίδευσης που θα προκύψει εφαρμόστε την μέθοδο *PCA* ώστε να μειώσετε τη διάστασή των δεδομένων. Δοκιμάστε διάφορες τιμές για τον αριθμό της μειωμένης διάστασης $M = \{100\%, 75\%, 50\%, 25\%, 10\%\}$ της πραγματικής διάστασης των δεδομένων. Στη συνέχεια, μειώστε την διάσταση και στα δεδομένα του συνόλου ελέγχου (με βάση τα ιδιοδιανύσματα που βρήκατε στο σύνολο εκπαίδευσης).

Κατασκευάστε έναν ταξινομητή προσώπων χρησιμοποιώντας την απλή μέθοδο των k κοντινότερων γειτόνων (k -NN). Χρησιμοποιήστε δύο είδη αποστάσεων: α) Ευκλείδεια απόσταση και β) συνημίτονο ($1 - \cos(x, y)$).

(β Ομαδοποίηση Προσώπων): Αφού πρώτα εφαρμόσετε την μέθοδο PCA σε όλες τις εικόνες (400) δοκιμάζοντας όπως και προηγουμένως διάφορες τιμές του αριθμού M της μειωμένης διάστασης, $M = \{100\%, 75\%, 50\%, 25\%, 10\%\}$, κάντε ομαδοποίηση σε 10 ομάδες με τις μεθόδους *k-means* και *GMMs*. Σε κάθε περίπτωση, αξιολογήστε τα αποτελέσματα με την ποσότητα **Purity**: η κατηγορία κάθε ομάδας (c_j) καθορίζεται, μετά το τέλος της ομαδοποίησης, από την πλειοψηφούσα πραγματική κατηγορία (ω_k) μεταξύ των μελών της ομάδας. Τότε η ακρίβεια (**purity**) του παραπάνω καθορισμού υπολογίζεται μετρώντας το μέσο των σωστά ταξινομημένων σημείων. Δηλ.

$$\text{purity}(\Omega, \mathbb{C}) = \frac{1}{N} \sum_k \max_j |\omega_k \cap c_j|$$

Παραδώστε, τόσο τον κώδικα σε matlab, όσο και τις παρατηρήσεις σας γραπτώς.