

# ΕΡΓΑΣΙΑ ΣΤΙΣ ΠΟΛΥΜΕΣΙΚΕΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Εαρινό Εξάμηνο 2020-2021

Στόχος της εργασίας είναι η ανάπτυξη και παρουσίαση μιας απλής εφαρμογής σε μία βάση δεδομένων εικόνων ή ήχων ή video στην οποία θα γίνεται αναζήτηση με βάση το περιεχόμενο (image/sound/video retrieval system). Τα αντικείμενα θα εισάγονται και θα επεξεργάζονται στην βάση ως αριθμητικά διανύσματα (descriptor vectors) βασικών τους χαρακτηριστικών που θα επιλέξετε εσείς (π.χ. χρώμα, υφή, σχήμα, ένταση, κλπ.), ενώ θα υπάρχουν σύνδεσμοι αναφοράς και στα αρχεία των κανονικών αντικειμένων ώστε ο χρήστης να μπορεί (αν θέλει) να εμφανίσει ένα αντικείμενο και να το δει ή να το ακούσει. Η αναζήτηση της ομοιότητας μεταξύ δύο αντικειμένων θα βασίζεται στα διανύσματα αυτά και στην βαθμολόγηση της ομοιότητάς τους. Για την υλοποίηση του συστήματός σας μπορείτε να χρησιμοποιήσετε ένα σύστημα διαχείρισης βάσεων δεδομένων της επιλογής σας (π.χ. PostgreSQL), ενώ για την εφαρμογή μπορείτε να χρησιμοποιήσετε οποιαδήποτε έτοιμη βιβλιοθήκη, τεχνολογία, γλώσσα. Ως ενδεικτικά δεδομένα μπορείτε να χρησιμοποιήσετε έτοιμα δεδομένα που υπάρχουν διαθέσιμα στο διαδίκτυο (π.χ. <http://www.imageclef.org/wikidata>, <http://www.robots.ox.ac.uk/~vgg/data/oxbuildings/>, κ.α.), ή ακόμα και να δημιουργήσετε μία δική σας μικρή συλλογή.

Πιο συγκεκριμένα οι απαιτήσεις είναι οι εξής:

1. Κάντε μία σύντομη περιγραφή της εφαρμογής σας και των δυνατοτήτων της.
2. Για κάθε αντικείμενο, θα πρέπει να εισάγεται στην βάση δεδομένων το αντίστοιχο διάνυσμα ή τα αντίστοιχα διανύσματα (descriptor vectors) των χαρακτηριστικών που έχετε επιλέξει (π.χ. για εικόνες το texture, color, shape ή και συνδυασμούς αυτών). Υπάρχουν αρκετά πρότυπα που παράγουν τέτοια διανύσματα (όπως π.χ. για εικόνες: CIME, TLEP, SURF, CEDD, SIFT, GIST, κλπ. – δείτε την διαφάνεια 35 που αναφέρθηκε και στο μάθημα [mmdb-bycontent.pdf]). Μπορείτε φυσικά να χρησιμοποιήσετε έτοιμα διανύσματα που έχουν ήδη παραχθεί από σύνολα δεδομένων του διαδικτύου.
3. Τα αρχεία των πραγματικών αντικειμένων θα πρέπει να είναι αποθηκευμένα στην δευτερεύουσα μνήμη, ενώ θα πρέπει κατά την εισαγωγή των διανυσμάτων τους στην βάση δεδομένων να αποθηκεύεται και ένας σύνδεσμος (link), ώστε ένα αντικείμενο να μπορεί να διαβαστεί (εφόσον ζητηθεί) και να εμφανιστεί στον χρήστη.
4. Η εφαρμογή σας θα πρέπει να υποστηρίζει αναζήτηση ομοίων αντικειμένων με βάση τα χαρακτηριστικά που επιλέξατε (π.χ. για εικόνες: texture, color, shape) ή και συνδυασμούς αυτών, ανάλογα με τα αντίστοιχα διανύσματα που θα χρησιμοποιήσετε. Για την υποστήριξη της αναζήτησης, ο χρήστης θα δίνει ένα αντικείμενο στην αντίστοιχη διανυσματική μορφή (το οποίο δεν είναι απαραίτητο να βρίσκεται στη βάση δεδομένων), ενώ θα δηλώνει και έναν αριθμό k για το πλήθος των ομοίων αντικειμένων που επιθυμεί ως απάντηση (ερώτημα ομοιότητας k-NN). Εκτός από τα διανύσματα των k πιο ομοίων αντικειμένων θα πρέπει να εμφανίζονται και οι αντίστοιχοι σύνδεσμοι προς τα πραγματικά αντικείμενα ώστε ο χρήστης να μπορεί να τα δει (ή να τα ακούσει) αν το επιθυμεί.

5. Στην εφαρμογή σας η μέτρηση και η βαθμολόγηση της ομοιότητας δύο αντικειμένων θα γίνεται εφαρμόζοντας μία **συνάρτηση βαθμολόγησης πάνω στα descriptors των αντικειμένων**. Σε περίπτωση που χρησιμοποιήσετε δύο ή περισσότερα descriptors για ένα αντικείμενο, η μέτρηση και βαθμολόγηση της ομοιότητας μπορεί να γίνει εφαρμόζοντας έναν ενιαίο descriptor (ενώνοντας τις διαστάσεις από όλα τα διανύσματα των descriptors).

6. Για τις αποστάσεις των histograms ή των descriptors (μέσα από τις οποίες θα προκύψει και η συνάρτηση βαθμολόγησης της ομοιότητας) θα πρέπει να επιλέξετε τουλάχιστον **πέντε διαφορετικά μέτρα απόστασης** από τους πίνακες που παρουσιάζονται στις διαφάνειες 36-45 που αναφέρθηκαν στο μάθημα [mmdb-bycontent.pdf], ή ακόμα και άλλα μέτρα από τη βιβλιογραφία. Ο χρήστης στις ρυθμίσεις της εφαρμογής θα πρέπει να μπορεί να επιλέγει ένα συγκεκριμένο μέτρο για να εκτελεί τα ερωτήματά του.

7. Στη βάση δεδομένων θα πρέπει να υπάρχουν ενδεικτικά τουλάχιστον **200 αντικείμενα** τα οποία είτε μπορείτε να επιλέξετε εσείς ελεύθερα είτε να χρησιμοποιήσετε υποσύνολα ή και ολόκληρα έτοιμα σύνολα δεδομένων του διαδικτύου.

8. **Εκτελέστε ερωτήματα αναζήτησης ομοιότητας k-NN** για  $k=5, 10, 20$  δίνοντας ως είσοδο ένα αντικείμενο της επιλογής σας και εμφανίστε τα αποτελέσματα (χωρίς τα πραγματικά αντικείμενα) σε συνδυασμό με τα πέντε μέτρα απόστασης που υλοποιήσατε. Να συγκρίνετε οπτικά (ή ακουστικά) τα αποτελέσματα από τις πέντε διαφορετικές αποστάσεις και να αναφέρετε την αποτελεσματικότητα του συστήματός σας (μετρώντας το precision σε ποσοστό %). Στην περίπτωση  $k=20$  να εμφανίσετε και τα πραγματικά αντικείμενα στα αποτελέσματα (ονόματα αρχείων και στην περίπτωση εικόνων τις ίδιες τις 20 εικόνες σε μικρή κλίμακα). Επίσης, να αναφέρετε ποιο μέτρο ήταν το πιο αποτελεσματικό.

9. Προαιρετικά μπορείτε να διαμορφώσετε το **περιβάλλον διαχείρισης της εφαρμογής σας** ώστε να μπορεί να εκτελεστεί μέσω ενός browser (π.χ. php στον MapServer). Επίσης μπορείτε **προαιρετικά να χρησιμοποιήσετε κάποιο κατάλληλο index για την αποθήκευση των αντικειμένων στην βάση δεδομένων** ώστε να επιταχύνετε την εκτέλεση των k-NN ερωτημάτων.

Το κείμενο της εργασίας θα περιλαμβάνει **εξώφυλλο με τα στοιχεία του μαθήματος, τίτλο της εργασίας, τα στοιχεία σας, περιεχόμενα, ενότητες στις οποίες θα περιγράφονται τα παραπάνω ζητούμενα, σχήματα και στιγμιότυπα οθόνης (screenshots) όπου είναι απαραίτητο, οδηγίες και παραδείγματα εκτέλεσης, σχετική βιβλιογραφία και αναφορές**. Τα τμήματα του κώδικα γλώσσας θα βρίσκονται σε **παράρτημα**. Η εργασία θα παραδοθεί σε ηλεκτρονική μορφή (κείμενο σε Word ή PDF, σύνολο δεδομένων, λοιπά αρχεία, κώδικας), το αργότερο μέχρι την **Παρασκευή 11 Ιουνίου 2021** στο e-mail του διδάσκοντα.