

Στατιστική Επεξεργασία Σήματος και Μάθηση

2η Εργαστηριακή Άσκηση

Ακαδημαϊκό Έτος 2023/24

Ανάκτηση εικόνας από βάση δεδομένων

Στόχος αυτού του θέματος καθίσταται η εξοικείωση σας με τη διαδικασία της ανάκτησης εικόνων καθώς και με γραμμικές και μη γραμμικές τεχνικές μείωσης της διάστασης δεδομένων (linear and non-linear dimensionality reduction). Για τον σκοπό αυτό, σας δίνεται μια βάση εικόνων (αρχείο: **Database.rar**), που είναι ουσιαστικά ένας φάκελος με έναν αριθμό από εικόνες, από τις οποίες ορισμένες μοιάζουν μεταξύ τους και θα πρέπει να αναπτύξετε κατάλληλες ρουτίνες για την ανάκτησή τους εφαρμόζοντας τεχνικές μείωσης διάστασης δεδομένων.

Λαμβάνοντάς υπόψιν ότι εικόνες είναι σήματα με πολύ μεγάλη διάσταση, ο βασικός σκοπός των τεχνικών μείωσης διάστασης είναι να μειώσουν αυτόν τον πολυδιάστατο χώρο και να αναπαραστήσουν τα σήματα σε έναν χώρο μικρότερης διάστασης, επιχειρώντας να διατηρήσουν την δομή και την πληροφορία των αρχικών σημάτων. Αυτός ο χώρος δύναται να έχει πολύ μικρότερη διάσταση, ειδικά σε περιπτώσεις που υπάρχει υψηλή συσχέτιση μεταξύ των δεδομένων. Οι εν λόγω τεχνικές χωρίζονται σε δύο βασικές κατηγορίες: στις γραμμικές μεθόδους και στις μη γραμμικές μεθόδους μείωσης διάστασης.

Αναλυτικότερα, στα πλαίσια της εν λόγω άσκησης καλείστε να υλοποιήσετε ένα σύστημα ανάκτησης εικόνων χρησιμοποιώντας μια γραμμική μέθοδο μείωσης διάστασης την Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis - **PCA**) καθώς και ένα δεύτερο σύστημα ανάκτησης εικόνων που βασίζεται στους Αυτοκωδικοποιητές (**Autoencoders**), οι οποίοι αποτελούν μία μη γραμμική μέθοδο μείωσης διάστασης. Σχετικά με την PCA, ο βασικός πυρήνας της είναι ουσιαστικά μια διάσπαση ιδιοτιμών και επιμέρους συνιστωσών που ονομάζονται ιδιο-εικόνες. Από την άλλη, οι Autoencoders είναι τεχνητά νευρωνικά δίκτυα τα οποία προβάλλουν τα δεδομένα της εισόδου σε έναν χώρο χαμηλότερης διάστασης και

έπειτα ανακατασκευάζουν στην έξοδο τους την είσοδο. Σύμφωνα με την εικόνα 1, βασίζονται σε μία αρχιτεκτονική κωδικοποιητή- αποκωδικοποιητή (Encoder - Decoder), όπου ο encoder κωδικοποιεί τα δεδομένα υψηλής διάστασης σε ένα χώρο χαμηλής διάστασης (ονομάζεται latent space) και ο decoder παίρνει αυτά τα δεδομένα χαμηλής διάστασης και προσπαθεί να ανακατασκευάσει τα αρχικά πολυδιάστατα δεδομένα.

Στο φάκελο **test** δίνονται 10 εικόνες με τις οποίες θα πειραματιστείτε για τον έλεγχο του των συστημάτων ανάκτησης που θα κατασκευάσετε με την χρήση της PCA και των Autoencoders. Δίνοντας μία εικόνα από τον φάκελο test ως είσοδο σε αυτό το σύστημα θα πρέπει αυτό να την εντοπίζει στην βάση δεδομένων.

Το αρχείο Database.rar δίνεται στον ακόλουθο σύνδεσμο [drive](#)

Ζητούμενα:

α) Σύστημα ανάκτησης “PCA”: Κατασκευάστε ένα σύστημα ανάκτησης εικόνων μέσω της PCA. Χρησιμοποιήστε τις εικόνες στο αρχείο **database.rar** ως training σετ ώστε να υπολογίστε την PCA των εικόνων της συγκεκριμένης βάσης. Εν συνεχεία, με βάση τις εικόνες στο test φάκελο εξάγετε ποσοστά επιτυχούς ανάκτησης. Η επιτυχής ανάκτηση αναφέρεται στην εύρεση της καλύτερης δυνατής προσέγγισης, δηλαδή της εγγύτερης στην δοθείσα βάση εικόνων. Για να εξάγεται τα ποσοστά επιτυχούς ανάκτησης πειραματιστείτε με την διάσταση του χώρου όπου προβάλλονται οι εικόνες διατηρώντας τις 100, 50 και 10 κύριες συνιστώσες για κάθε εικόνα. Είναι σημαντικό για κάθε εικόνα να αναγνωρίζονται οι κύριες συνιστώσες για λόγους υπολογιστικής πολυπλοκότητας. Σχολιάστε τα αποτελέσματα.

β) Σύστημα ανάκτησης “Autoencoders”: Κατασκευάστε ένα σύστημα ανάκτησης εικόνων με την χρήση των Autoencoders. Όπως και στο προηγούμενο ερώτημα, χρησιμοποιήστε τις εικόνες στο αρχείο **database.rar** ώστε να εκπαιδεύσετε το δίκτυο σας και να προβάλλετε τις εικόνες της βάσης στον χώρο χαμηλότερης διάστασης μέσω του encoder δικτύου. Εν συνεχεία, χρησιμοποιήστε τις εικόνες στον test φάκελο και αφού τις προβάλλετε και αυτές στο χώρο χαμηλότερης διάστασης μέσω του encoder, εξάγετε ποσοστά επιτυχούς ανάκτησης. Για την εξαγωγή των ποσοστών επιτυχούς ανάκτησης, πειραματιστείτε όπως και πριν με την διάσταση του χώρου προβάλλοντας τις εικόνες σε ένα χώρο με διάσταση 100, 50 και 10. Ουσιαστικά θα πρέπει να εκπαιδεύσετε 3 ξεχωριστούς autoencoders, έναν για κάθε διάσταση που ζητείται.

Σχετικά με την υλοποίηση του συστήματος, κατασκευάστε autoencoders οι οποίοι αποτελούνται από ένα layer encoder και ένα layer decoder. Επομένως, η μαθηματική σχέση που περιγράφει τον autoencoder δίνεται από τον ακόλουθο τύπο:

$$\hat{x} = \varphi_{\text{dec}}(W_{\text{dec}} * (\varphi_{\text{enc}}(W_{\text{enc}} * x + b_{\text{enc}})) + b_{\text{dec}}) \quad (1)$$

Όπου W_{dec} είναι τα βάρη του decoder layer, b_{dec} είναι ο αντίστοιχος bias όρος, W_{enc} , b_{enc} είναι τα βάρη και ο αντίστοιχος bias όρος του encoder layer. Επίσης φ_{dec} και φ_{enc} είναι η συνάρτηση ενεργοποίησης του decoder και encoder layer αντίστοιχα.

1. Σε αυτό το ερώτημα σχεδιάστε έναν απλοποιημένο autoencoder. Αφαιρέστε τις συναρτήσεις ενεργοποίησης και τους bias όρους από τα encoder και decoder layers και θεωρείστε την ακόλουθη μαθηματική σχέση

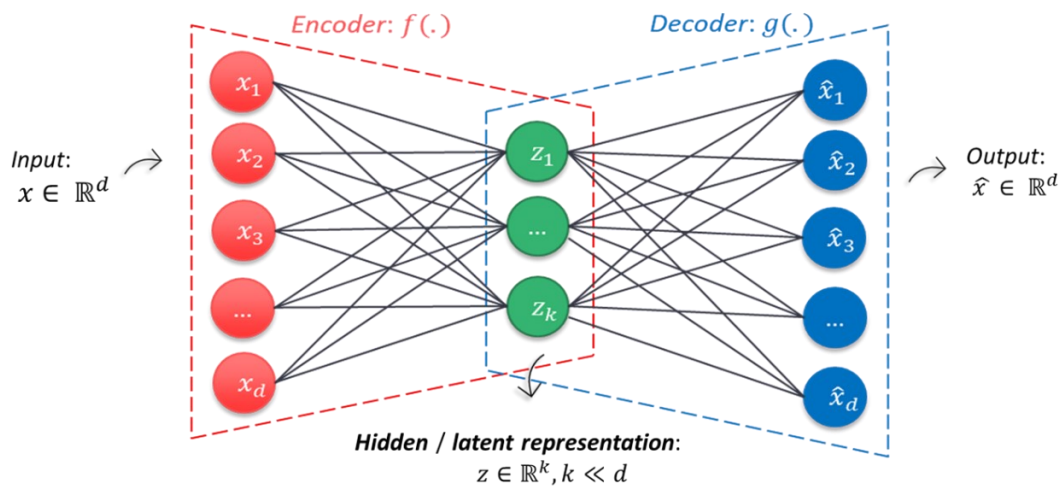
$$\hat{x} = W_{\text{dec}}(W_{\text{enc}} * x)$$

$$\text{Όπου } W_{\text{enc}} = W \text{ και } W_{\text{dec}} = W^T$$

Σχολιάστε και συγκρίνεται τα αποτελέσματα με το προηγούμενο ερώτημα. Ποια η σχέση αυτής της μαθηματικής σχέσης εν συγκρίσει με την προσέγγιση του προβλήματος με την PCA.

2. Επαναλάβετε το ερώτημα β. Σε αυτό το ερώτημα, χρησιμοποιήστε έναν κανονικό autoencoder όπως περιγράφεται από την σχέση (1). Πειραματιστείτε με την συνάρτηση ενεργοποίησης.

Σχολιάστε και συγκρίνεται τα αποτελέσματα με τα προηγούμενα ερώτημα.



Εικόνα 1 Δομή ενός autoencoder.

Βιβλιογραφία

- [1]. Lindsay I Smith, "A tutorial on Principal Components Analysis", Cornell University, 2002.
- [2]. Jonathon Shlens, "A tutorial on Principal Components Analysis", arXiv preprint, 2014.
- [3]. Turk, M. and Pentland A., "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, 1991.
- [4]. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning, MIT Press, 2016, <http://www.deeplearningbook.org>.
- [5]. Van Der Maaten, L., Postma, E. & Van den Herik, J. (2009). Dimensionality reduction: a comparative review. J Mach Learn Res, 10, 66-71.
- [6]. David Charte, Francisco Charte, Salvador Garcia, Maria J. del Jesus, and Francisco Herrera, "A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines," Information Fusion, vol. 44, pp. 78–96, Nov 2018.