

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ ΚΑΙ ΒΑΘΙΑ ΜΑΘΗΣΗ

8ο εξάμηνο

Ακαδημαϊκό έτος 2022-2023

2η Σειρά Αναλυτ. Ασκήσεων

Ημερ. Παράδ.: 19.06.2023

Γενικές Οδηγίες: Οι αναλυτικές σειρές ασκήσεων είναι ατομικές, και οι λύσεις που θα δώσετε πρέπει να αντιπροσωπεύουν μόνο την προσωπική σας εργασία. Εξηγήστε επαρκώς την εργασία σας. Αν χρησιμοποιήσετε κάποια άλλη πηγή εκτός των σημειώσεων για την λύση σας, πρέπει να το αναφέρετε. Η παράδοση των λύσεων των αναλυτικών ασκήσεων της σειράς αυτής θα γίνει ηλεκτρονικά στην HELIOS ιστοσελίδα του μαθήματος και θα πρέπει να την υποβάλετε ως ένα ενιαίο αρχείο PDF με το εξής filename format χρησιμοποιώντας μόνο λατινικούς χαρακτήρες: DL23_hwk2_AM_LastnameFirstname.pdf, όπου AM είναι ο 8-ψήφιος αριθμός μητρώου σας. Σκαναριμένες χειρόγραφες λύσεις επιτρέπονται αρκεί να είναι καθαρογραμμένες και ευανάγνωστες. Επίσης στην 1η σελίδα των λύσεων θα αναγράφετε το ονοματεπώνυμο, Α.Μ., και το email σας. Να σημειωθεί ότι η καταληκτική ημερομηνία παράδοσης είναι τελική και δεν θα δοθεί παράταση.

Άσκηση 2.1 (Attention - Transformers)

1. Έστω η εξής ακολουθία μήκους 4:

$$x^{(1)} = [-2.0, 1.0, 0.5]^T, \quad x^{(2)} = [1.0, 1.5, -0.5]^T, \quad x^{(3)} = [-1.5, 1.0, -0.5]^T, \quad x^{(4)} = [-2.0, -2.5, 1.5]^T$$

, της οποίας κάθε στοιχείο είναι ένα διάνυσμα (vector) στο \mathbb{R}^3 .

Έστω επίσης ότι με βάση τα παραπάνω διανύσματα έχουμε έναν πίνακα εισόδου (input matrix) $X = [x^{(1)}, x^{(2)}, x^{(3)}, x^{(4)}]^T \in \mathbb{R}^{4 \times 3}$. Έστω ότι θέλουμε να χρησιμοποιήσουμε ένα επίπεδο προσοχής (attention layer) στη συγκεκριμένη είσοδο με query vector $q = [-2.0, 1.0, -1.0]^T$. Υπολογίστε τις πιθανότητες προσοχής (attention probabilities) και το διάνυσμα εξόδου (output vector) με βάση το query X και κάνοντας χρήση scaled dot product attention.

2. Ας υποθέσουμε ότι αντί για τη χρήση ενός query vector θέλουμε να υπολογίσουμε (single-head) self-attention για την ίδια είσοδο. Έστω ότι τα projection matrices για τα queries, keys, και values είναι αντίστοιχα:

$$W_Q = \begin{bmatrix} 1 & -1.5 \\ 0 & 2 \\ -0.5 & -1 \end{bmatrix}, \quad W_K = \begin{bmatrix} -1.5 & -1 \\ 2.5 & 0 \\ 0.5 & -1 \end{bmatrix}, \quad W_V = \begin{bmatrix} 1 & 2.5 \\ -0.5 & -2 \\ 0 & -1 \end{bmatrix},$$

Υπολογίστε τους πίνακες query, key, και value (Q, K, V), καθώς και το διάνυσμα εξόδου Z.

Άσκηση 2.2 (Generative Models)

Έστω ότι θέλετε να δημιουργήσετε ρεαλιστικές εικόνες χταποδιών χρησιμοποιώντας Generative Adversarial Networks (GANs). Ωστόσο δυστυχώς δεν έχετε καταφέρει να συγκεντρώσετε μεγάλο αριθμό φωτογραφιών από αληθινά χταπόδια, και γι' αυτό αποφασίζετε να χρησιμοποιήσετε τεχνικές επαύξησης δεδομένων (data augmentation) με την ελπίδα ότι με αυτόν τον τρόπο θα βελτιώσετε την ποιότητα της εκπαίδευσης του δικτύου σας. Δοκιμάζετε τρεις δημοφιλείς τεχνικές επαύξησης:

1. Θόλωμα εικόνων
2. Αλλαγή χρώματος των pixels
3. Περιστροφή (flip) ως προς τον κεντρικό κατακόρυφο άξονα της εικόνας (άξονας "αριστερά-δεξιά")

Ποια/-ες από τις παραπάνω τεχνικές θεωρείτε ότι θα ήταν κατάλληλη/-ες για να βελτιώσει/-ουν την ποιότητα των εικόνων που παράγει ο Generator; Αιτιολογήστε αναλυτικά την απάντησή σας. Εάν μπορείτε, προτείνετε κάποια επιπλέον τεχνική που θεωρείτε ότι θα ήταν αποτελεσματική για τον ίδιο σκοπό.

Άσκηση 2.3 (Self-Supervised Learning)

Εξηγήστε τις βασικές αρχές της αυτοεπιβλεπόμενης αντιθετικής μάθησης (self-supervised contrastive learning). Μετά από κατάλληλη αναζήτηση και μελέτη στην πρόσφατη βιβλιογραφία, περιγράψτε και συγκρίνετε δύο τεχνικές αυτοεπιβλεπόμενης αντιθετικής μάθησης με εφαρμογή στο πεδίο της ανάλυσης εικόνας / όρασης υπολογιστών και δύο (διαφορετικές) τεχνικές με εφαρμογή στο πεδίο της γλωσσικής μοντελοποίησης / επεξεργασίας φυσικής γλώσσας.