



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ & ΒΑΘΙΑ ΜΑΘΗΣΗ

2^η ΣΕΙΡΑ ΓΡΑΠΤΩΝ ΑΣΚΗΣΕΩΝ

Ειρήνη Δόντη

ΑΜ: 03119839

8ο εξάμηνο

Αθήνα 2023

Άσκηση 2.1

1. Κάνουμε χρήση της scaled dot product attention.

Από την εκφώνηση, ισχύουν τα εξής:

$$dk = 3, h = X = \begin{bmatrix} -2.0 & 1.0 & 0.5 \\ 1.0 & 1.5 & -0.5 \\ -1.5 & 1.0 & -0.5 \\ -2.0 & -2.5 & 1.5 \end{bmatrix} \text{ και } q = [-2.0 \ 1.0 \ -1.0]^T$$

Υπολογίζουμε τις πιθανότητες προσοχής: $p = \text{softmax}(S) = \text{softmax}(\frac{Xq}{\sqrt{dk}}) =$

$$\text{softmax}(\frac{1}{\sqrt{3}} \begin{bmatrix} -2.0 & 1.0 & 0.5 \\ 1.0 & 1.5 & -0.5 \\ -1.5 & 1.0 & -0.5 \\ -2.0 & -2.5 & 1.5 \end{bmatrix} [-2.0 \ 1.0 \ -1.0]^T) =$$

$$\text{softmax}(\begin{bmatrix} \frac{4.5}{\sqrt{3}} & 0 & \frac{4.5}{\sqrt{3}} & 0 \end{bmatrix}^T) = \text{softmax}([2.598 \ 0 \ 2.598 \ 0]^T) =$$

$$[0.465 \ 0.035 \ 0.465 \ 0.035]^T$$

Υπολογίζουμε το διάνυσμα εξόδου: $Z = h^T p = h^T \text{softmax}(\frac{Xq}{\sqrt{dk}}) =$

$$\begin{bmatrix} -2.0 & 1.0 & -1.5 & -2.0 \\ 1.0 & 1.5 & 1.0 & -2.5 \\ 0.5 & -0.5 & -0.5 & 1.5 \end{bmatrix} [0.465 \ 0.035 \ 0.465 \ 0.035]^T =$$

$$[-1.663 \ 0.895 \ 0.035]^T$$

2. Ισχύει το εξής: $Q = XW^Q, K = XW^K, V = XW^V$

Οπότε, οι πίνακες Q, K, V είναι ίσοι με:

$$Q = \begin{bmatrix} -2.0 & 1.0 & 0.5 \\ 1.0 & 1.5 & -0.5 \\ -1.5 & 1.0 & -0.5 \\ -2.0 & -2.5 & 1.5 \end{bmatrix} \begin{bmatrix} 1 & -1.5 \\ 0 & 2 \\ -0.5 & -1 \end{bmatrix} = \begin{bmatrix} -2.25 & 4.5 \\ 1.25 & 2 \\ -1.25 & 4.75 \\ -2.75 & -3.5 \end{bmatrix}$$

$$K = \begin{bmatrix} -2.0 & 1.0 & 0.5 \\ 1.0 & 1.5 & -0.5 \\ -1.5 & 1.0 & -0.5 \\ -2.0 & -2.5 & 1.5 \end{bmatrix} \begin{bmatrix} -1.5 & -1 \\ 2.5 & 0 \\ 0.5 & -1 \end{bmatrix} = \begin{bmatrix} 5.75 & 1.5 \\ 2 & -0.5 \\ 4.5 & 2 \\ -2.5 & 0.5 \end{bmatrix}$$

$$V = \begin{bmatrix} -2.0 & 1.0 & 0.5 \\ 1.0 & 1.5 & -0.5 \\ -1.5 & 1.0 & -0.5 \\ -2.0 & -2.5 & 1.5 \end{bmatrix} \begin{bmatrix} 1 & 2.5 \\ -0.5 & -2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} -2.5 & -7.5 \\ 0.25 & 0 \\ -2 & -5.25 \\ -0.75 & -1.5 \end{bmatrix}$$

Το διάνυσμα εξόδου είναι: $Z = \text{softmax}(\frac{QK^T}{\sqrt{dk}})V =$

$$\begin{aligned} & \text{softmax}(\frac{1}{\sqrt{dk}} \begin{bmatrix} -2.25 & 4.5 \\ 1.25 & 2 \\ -1.25 & 4.75 \\ -2.75 & -3.5 \end{bmatrix} \begin{bmatrix} 5.75 & 2 & 4.5 & -2.5 \\ 1.5 & -0.5 & 2 & 0.5 \end{bmatrix}) \begin{bmatrix} -2.5 & -7.5 \\ 0.25 & 0 \\ -2 & -5.25 \\ -0.75 & -1.5 \end{bmatrix} = \\ & \text{softmax}(\frac{1}{\sqrt{2}} \begin{bmatrix} -6.19 & -6.75 & -1.125 & 7.88 \\ 10.19 & 1.50 & 9.625 & -2.13 \\ -0.06 & -4.88 & 3.88 & 5.50 \\ -21.06 & -3.75 & -19.38 & 5.12 \end{bmatrix}) \begin{bmatrix} -2.5 & -7.5 \\ 0.25 & 0 \\ -2 & -5.25 \\ -0.75 & -1.5 \end{bmatrix} = \\ & = \text{softmax}(\begin{bmatrix} -4.38 & -4.77 & -0.795 & 5.57 \\ 7.21 & 1.06 & 6.81 & -1.51 \\ -0.04 & -3.45 & 2.74 & 3.89 \\ -14.89 & -2.65 & -13.70 & 3.62 \end{bmatrix}) \begin{bmatrix} -2.5 & -7.5 \\ 0.25 & 0 \\ -2 & -5.25 \\ -0.75 & -1.5 \end{bmatrix} = \\ & \begin{bmatrix} 4.76 \times 10^{-5} & 3.23 \times 10^{-5} & 0.0017 & 0.998 \\ 0.598 & 0.0013 & 0.401 & 9.76 \times 10^{-5} \\ 0.0147 & 4.855 \times 10^{-4} & 0.237 & 0.748 \\ 9.13 \times 10^{-9} & 0.002 & 3 \times 10^{-8} & 0.998 \end{bmatrix} \begin{bmatrix} -2.5 & -7.5 \\ 0.25 & 0 \\ -2 & -5.25 \\ -0.75 & -1.5 \end{bmatrix} = \\ & \begin{bmatrix} -0.752 & -1.506 \\ -2.297 & -6.590 \\ -1.072 & -2.476 \\ -0.748 & -1.497 \end{bmatrix} \end{aligned}$$

Άσκηση 2.2

Για τη δημιουργία ρεαλιστικών εικόνων χταποδιών χρησιμοποιώντας Generative Adversarial Networks (GANs) και επαύξηση δεδομένων, μπορούμε να αξιοποιήσουμε τις παραπάνω τεχνικές ως εξής:

1. Θόλωμα Εικόνων

Το θόλωμα μπορεί να βοηθήσει στην παραγωγή πιο ομαλών και απαλών σχημάτων στις εικόνες. Συνεπώς, οδηγεί στην πιο φυσική εμφάνιση των χταποδιών και μειώνει τον αριθμό των αιχμηρών ή ακραίων λεπτομερειών που μπορεί να δημιουργήσει το δίκτυο. Με άλλα λόγια, παράγονται πιο ρεαλιστικές εικόνες χταποδιών.

2. Αλλαγή χρώματος των pixels

Η αλλαγή του χρώματος των pixels μπορεί να δημιουργήσει ποικιλία στις εικόνες και να επιτρέψει στον Generator να υιοθετήσει ποικιλία χρωμάτων και αποχρώσεων που σχετίζονται με τα χταπόδια. Εφαρμόζοντας τυχαίες μικρές αλλαγές στο χρώμα των pixels, όπως αλλαγές στην ένταση, τον κορεσμό ή την απόχρωση, δημιουργείται μια ποικιλία από εικόνες χταποδιών.

3. Περιστροφή (flip) ως προς τον κεντρικό κατακόρυφο άξονα της εικόνας

Η περιστροφή των εικόνων μπορεί να βοηθήσει τον Generator να υιοθετήσει την αναγνώριση και την απεικόνιση των χταποδιών από διάφορες γωνίες. Με την περιστροφή, μπορεί να δημιουργηθούν περισσότερες εκδοχές των εικόνων χταποδιών, λαμβάνοντας υπόψη διάφορες οπτικές γωνίες.

Επιπλέον τεχνική:

4. Προσθήκη θορύβου:

Εισάγοντας τυχαίο θόρυβο στις εικόνες χταποδιών, αυξάνουμε την ποικιλία των δεδομένων εκπαίδευσης. Ο θόρυβος μπορεί να προσομοιώσει διάφορες ατέλειες ή αβεβαιότητες που μπορεί να συναντώνται στις πραγματικές εικόνες χταποδιών, όπως θόλωμα, θόρυβος κάμερας ή παραμόρφωση.

Επιλέγοντας τις παραπάνω τεχνικές, δημιουργείται μια ποικιλία από εικόνες χταποδιών με διαφορετικές γωνίες, χρώματα και ατέλειες, βοηθώντας τον Generator να μάθει τα χαρακτηριστικά και την ποικιλία των εικόνων χταποδιών. Είναι σημαντικό να θυμόμαστε ότι η επαύξηση δεδομένων πρέπει να γίνεται με ισορροπημένο τρόπο, ώστε να μην παραμορφώνονται υπερβολικά οι εικόνες και να χάνουν την αρχική τους πληροφορία. Επίσης, η επιλογή των τεχνικών επαύξησης δεδομένων εξαρτάται επίσης από τη διαθεσιμότητα δεδομένων και τις ειδικές απαιτήσεις του προβλήματος.

Άσκηση 2.3

Οι βασικές αρχές της αυτοεπιβλεπόμενης αντιθετικής μάθησης (self-supervised contrastive learning) είναι η χρήση ενός αυτοεπιβλεπόμενου προβλήματος για την εκπαίδευση ενός μοντέλου μη επιβλεπόμενης μάθησης. Οι αυτοεπιβλεπόμενες αλγόριθμοι αντιθετικής μάθησης αναπτύχθηκαν, για να εξάγουν χρήσιμες αναπαραστάσεις από μη επιβλεπόμενα δεδομένα, χρησιμοποιώντας την αντιθετική και συγκριτική πληροφορία.

Οι **βασικές αρχές της αυτοεπιβλεπόμενης αντιθετικής μάθησης** περιλαμβάνουν:

Επιλογή ενός αυτοεπιβλεπόμενου προβλήματος: Ορισμός μιας εργασίας που απαιτεί το μοντέλο να εξάγει αναπαραστάσεις από τα δεδομένα χωρίς να υπάρχουν ετικέτες κλάσεων.

Αντιθετική αναπαράσταση: Οι μέθοδοι αυτοεπιβλεπόμενης αντιθετικής μάθησης επιδιώκουν να μάθουν αναπαραστάσεις όπου η ομοιότητα και η αντιθεσιμότητα μεταξύ των δειγμάτων αντιστοιχίζονται σε κοντινές και απομακρυσμένες αποστάσεις στον χώρο των αναπαραστάσεων αντίστοιχα.

Εκπαίδευση με αντίθεση: Κατά την εκπαίδευση, το μοντέλο προσπαθεί να μάθει να εξάγει αναπαραστάσεις που μπορούν να διακρίνουν το ίδιο δείγμα από διαφορετικές παραλλαγές του, ενώ ταυτόχρονα να απομακρύνουν τις αναπαραστάσεις διαφορετικών δειγμάτων.

Σχετικά με τις **τεχνικές αυτοεπιβλεπόμενης αντιθετικής μάθησης στο πεδίο της ανάλυσης εικόνας / όρασης υπολογιστών**, δύο δημοφιλείς τεχνικές είναι οι:

Αυτοκωδικοποιητές (Autoencoders)

Οι αυτοκωδικοποιητές αποτελούν ένα είδος αυτοεπιβλεπόμενων αντιθετικών μοντέλων. Οι αυτοκωδικοποιητές εκπαιδεύονται να ανακαλύπτουν την πληροφορία στην είσοδο και να παράγουν συμπίεσμένες αναπαραστάσεις της και στη συνέχεια να ανακατασκευάζουν την αρχική είσοδο από αυτές τις αναπαραστάσεις. Οι αυτοκωδικοποιητές επιδιώκουν να εξάγουν συμπίεσμένες αναπαραστάσεις που κωδικοποιούν την κύρια δομή και τα χαρακτηριστικά των εικόνων, ανεξάρτητα από τις ετικέτες κλάσης.

Μέθοδοι αυτοεπίβλεψης με χρήση επαγωγικών μοντέλων

Αυτή η τεχνική επιδιώκει να εκμεταλλευτεί τις επαγωγικές ιδιότητες των εικόνων. Επιτυγχάνεται αυτοεπίβλεψη μέσω της εισαγωγής παραλλαγών στις εικόνες, όπως αλλοιώσεις, περικοπές ή αναστροφές, και την εκπαίδευση ενός μοντέλου για να αναγνωρίσει αν οι δύο εικόνες προέρχονται από την ίδια αρχική εικόνα ή όχι.

Σχετικά με τις τεχνικές αυτοεπιβλεπόμενης αντιθετικής μάθησης στο πεδίο της γλωσσικής μοντελοποίησης / επεξεργασίας φυσικής γλώσσας, δύο διαφορετικές τεχνικές είναι:

Προβλέψεις λέξης (Word Predictions)

Σε αυτήν την τεχνική, ένα μοντέλο εκπαιδεύεται να προβλέπει την επόμενη λέξη σε μια ακολουθία λέξεων. Με την εκπαίδευση σε ένα μεγάλο σύνολο δεδομένων κειμένου, το μοντέλο μπορεί να μάθει τη σύνταξη, τη σημασιολογία και τις συσχετίσεις μεταξύ των λέξεων.

Πρόβλεψη πρότασης (Sentence Prediction)

Αυτή η τεχνική εκπαιδεύει ένα μοντέλο να προβλέπει αν μια πρόταση είναι έγκυρη ή όχι, χρησιμοποιώντας ένα αυτοδημιούργητο σύνολο δεδομένων. Η εκπαίδευση σε αυτό το πρόβλημα απαιτεί την κατανόηση των συντακτικών και σημασιολογικών χαρακτηριστικών των προτάσεων.

Συγκρίνοντας αυτές τις δύο τεχνικές, οι αυτοκωδικοποιητές εστιάζουν στην εξαγωγή αναπαραστάσεων από εικόνες, ενώ οι μέθοδοι επαγωγικών μοντέλων εστιάζουν στην εκμάθηση περίτεχνων χαρακτηριστικών των εικόνων. Ωστόσο, και οι δύο τεχνικές αποσκοπούν στην αντιμετώπιση του προβλήματος της μη επιβλεπόμενης μάθησης και της απόκτησης αναπαραστάσεων χωρίς την ανάγκη ετικετών κλάσεων.