

Προγραμματισμός για Γλωσσική Τεχνολογία II (M906)

Εργασίες



Τελική εργασία - Ήχος



Στοιχεία εργασίας

Τίτλος:

Τελική εργασία - Ήχος

Περιγραφή:

Σκοπός εργασίας

Σκοπός αυτής της εργασίας είναι να δημιουργήσετε και να αξιολογήσετε ένα σύστημα που αναγνωρίζει σε πραγματικό χρόνο συναίσθημα από ομιλία.

Η εργασία μπορεί να πραγματοποιηθεί ατομικά ή σε ομάδες 2 ατόμων.

Αρχείο καταγραφής/αναζήτησης αποριών:

https://imisathena-my.sharepoint.com/:w:/g/personal/maximos_athenarc_gr/EW9pqP1pzw5DIgWu6H4X1_IBXyOXAvf8gwwGTui65YJA7w?e=LrlsvX

Αρχεία αναφοράς

Σαν δεδομένα εκπαίδευσης μπορείτε να χρησιμοποιήσετε τα δεδομένα που έχουν ήδη υποστεί επεξεργασία σε προηγούμενο μάθημα (prepared_dataframe.pickle), στον φάκελο:

https://gitlab.com/language-technology-msc/programming-for-language-technology-ii-2021-2022/programminglangtechii_c07

Μπορείτε να εκπαιδεύσετε το μοντέλο μηχανικής μάθησης (της επιλογής σας) όπως κάναμε για τα τρία μοντέλα εδώ:

https://gitlab.com/language-technology-msc/programming-for-language-technology-ii-2021-2022/programminglangtechii_c07/-/blob/main/3_run_explore_data.py

Αν χρειαστείτε πρόσβαση στα αρχεία ήχου που χρησιμοποιήθηκαν για όλα τα παραπάνω, μπορείτε να τα βρείτε εδώ:

https://imisathena-my.sharepoint.com/:f:/g/personal/maximos_athenarc_gr/EnFYTi0o5bjMJYPo1IPbjkMB8w55AS43mBQjMuObjtF2Q?e=zeMyDb

Μπορείτε να χρησιμοποιήσετε τις υλοποιήσεις στο τελευταίο μάθημα για να λαμβάνεται κυλιώμενα τμήματα ήχου επιθυμητού μεγέθους σε πραγματικό χρόνο από αρχεία ήχου (σαν "προσομοίωση" εισόδου από μικρόφωνο):

https://gitlab.com/language-technology-msc/programming-for-language-technology-ii-2021-2022/programminglangtechii_c12/-/blob/main/emotion_real_time/1_file_test_realtime_processing.py

ή για λήψη ήχου από μικρόφωνο σε πραγματικό χρόνο:

https://gitlab.com/language-technology-msc/programming-for-language-technology-ii-2021-2022/programminglangtechii_c12/-/blob/main/emotion_real_time/2_mic_realtime_processing.py

Ζητούμενα:

1) Επιλέξτε ένα πρόβλημα κατηγοριοποίησης συναισθημάτων (π.χ. "calm"- "angry"), επιλέξτε τα κατάλληλα χαρακτηριστικά ήχου (ανάμεσα σε MFCCs, spectral centroid & bandwidth) και επιλέξτε και εκπαιδεύστε ένα μοντέλο μηχανικής μάθησης (όσο απλό κι αν είναι), παρουσιάζοντας αποτελέσματα της αξιολόγησής του, όχι σε πραγματικό χρόνο (δηλαδή με τη διαδικασία χωρισμού σε δεδομένα εκπαίδευσης-ελέγχου, όπως κάναμε στο μάθημα 8). Αποθηκεύστε το μοντέλο αυτό για φόρτωση και χρήση στο πρόγραμμα πραγματικού χρόνου που θα φτιάξετε αργότερα.

2) Αξιολογήστε το μοντέλο που αποθηκεύσατε παραπάνω σε λειτουργία πραγματικού χρόνου, αλλά με τον ήχο να προέρχεται από δεδομένα αρχείων ήχου (για έλεγχο σε "εργαστηριακές συνθήκες"). Είναι δική σας επιλογή το σε πόσα αρχεία θα το αξιολογήσετε και τι μετρική θα χρησιμοποιήσετε (π.χ. θα αξιολογείτε την αποτελεσματικότητα σε παράθυρα χρόνου ή θα αξιολογείτε στο σύνολο του αρχείου;).

3) Τρέξτε το μοντέλο με τη φωνή σας και τυπώστε τα αποτελέσματα όπως επιθυμείτε (π.χ. πάνω σε γράφημα ή ακόμα και ως τυπώματα στην κονσόλα). Αποθηκεύστε το αποτέλεσμα με όποια μορφή θεωρείτε επαρκώς περιγραφική (π.χ. screenshot ή αντιγραφή-επικόλληση από την κονσόλα).

4) Προσπαθείστε σύντομα να προτείνετε επόμενα πιθανά βήματα για διορθώσεις των αποτελεσμάτων (λαμβάνοντας υπόψη τη διαδικασία επεξεργασίας των δεδομένων, με την οποία δεν ήταν απαραίτητο να ασχοληθείτε ως τώρα).

Παραδοτέα:

1) Κώδικας που παρήγαγε τα παραπάνω (σε ένα ή πολλά ξεχωριστά αρχεία), ο οποίος τρέχει χωρίς σφάλματα.

2) Κείμενο με απλή περιγραφή για το ποια ηχητικά χαρακτηριστικά επιλέξατε, ποιο μοντέλο, πως το αξιολογήσατε σε πραγματικό χρόνο (στο Ζητούμενο 2), κείμενο/πίνακες αποτελεσμάτων και εικόνες που να δείχνουν συνοπτικά και περιεκτικά τα αποτελέσματα και κείμενο με σύντομα σχόλια για πιθανά βήματα διορθώσεων. Μπορείτε να το συμπεριλάβετε τα κείμενα αυτά στα αρχεία κώδικα, αν π.χ. το ετοιμάσετε με jupyter notebook ή αν βάλετε το σχετικό κείμενο με σχόλια στον κώδικα και στείλετε σε ξεχωριστό φάκελο τις όποιες εικόνες.

Μέγιστη βαθμολογία:

10

Τύπος Βαθμολογίας:

Αριθμός

Ημερομηνία έναρξης:

07-06-2022 09:27:00

Προθεσμία υποβολής:

10-07-2022 23:55:00

(απομένουν 18 ημέρες 8 ώρες 19 λεπτά)

Τύπος εργασίας:

Ομαδική εργασία

Η εργασία αυτή είναι ομαδική.

Θα πρέπει να εγγραφείτε σε κάποια ομάδα πριν μπορέσετε να υποβάλετε κάποιο αρχείο.

