

Θέματα Πτυχιακής/Μεταπτυχιακής Εργασίας

Εργαστήριο Επεξεργασίας Σήματος Φωνής

21 Φεβρουαρίου 2024

Τα παρακάτω θέματα παρουσιάζουν πτυχιακές/μεταπτυχιακές εργασίες στο Εργαστήριο Επεξεργασίας Σήματος Φωνής του Τμήματος Επιστήμης Υπολογιστών.

Απαραίτητη προϋπόθεση εκ μέρους του/της φοιτητή/τριας για την διεξαγωγή εργασίας αποτελούν τα παρακάτω:

- Βαθμός ≥ 8.5 στο μάθημα ΗΥ215-Εφαρμοσμένα Μαθηματικά για Μηχανικούς
- Βαθμός ≥ 8.5 στο μάθημα ΗΥ370-Ψηφιακή Επεξεργασία Σήματος
- (Προαιρετικό) Παρακολούθηση ή επιτυχία στο μάθημα ΗΥ578-Ψηφιακή Επεξεργασία Φωνής

- 1. Ερευνητικό: Προσαρμόσιμοι Mel-Frequency Cepstral Συντελεστές (aMFCCs) - 12 μήνες:**
Οι συντελεστές MFCCs είναι μια πολύ δημοφιλής αναπαράσταση σημάτων ήχου και ομιλίας, με εφαρμογές στην αναγνώριση ομιλίας και ομιλητή, καθώς και στην εξαγωγή χαρακτηριστικών της ομιλίας. Οι MFCCs εξάγονται από μια συγκεκριμένη ακολουθία βημάτων που βασίζεται στο Διακριτό Μετασχηματισμό Fourier (DFT), ο οποίος εφαρμόζεται σε μικρά τμήματα του σήματος. Ο τελευταίος υποθέτει ότι το τμήμα του σήματος είναι στάσιμο, δηλ. η συχνοτική του πληροφορία δεν αλλάζει σημαντικά μέσα σε αυτό το τμήμα. Όμως γνωρίζουμε πλέον ότι αυτή η υπόθεση δεν είναι πάντα αληθής. Στόχος της πτυχιακής είναι η μοντελοποίηση του φάσματος του σήματος με πιο προηγμένα μοντέλα ανάλυσης όπως το προσαρμόσιμο σχεδόν αρμονικό μοντέλο (eaQHM). Τέτοια μοντέλα έχουν υψηλότερη ανάλυση μέσα σε ένα τμήμα σήματος. Θέλουμε λοιπόν να παράξουμε MFCCs που βασίζονται σε ένα καλύτερο φάσμα που έχει παραχθεί από τέτοια, υψηλής ανάλυσης μοντέλα. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε MATLAB ή Python, καθώς και υπόβαθρο σε επεξεργασία σήματος φωνής.
- 2. Ερευνητικό: Ανάλυση και Αναγνώριση Συναισθήματος μέσω Σημάτων Φωνής με χρήση τεχνικών Μηχανικής Μάθησης - 12 μήνες:** Η αναγνώριση συναισθήματος αποτελεί σημαντικό κεφάλαιο στην Επεξεργασία Φωνής. Υπάρχουν πολλά χαρακτηριστικά (features) της ομιλίας που μπορούν να χρησιμοποιηθούν προς αυτήν την κατεύθυνση. Θα θέλαμε να χρησιμοποιήσουμε αυτά τα χαρακτηριστικά για την αναγνώριση συναισθήματος ομιλητών με τεχνικές μηχανικής μάθησης (Support Vector Machines, Random Forests, Ensemble Methods). Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση φωνής και των χαρακτηριστικών της, τεχνικών μηχανικής μάθησης, καθώς και γνώση χαρακτηριστικών της εκφραστικής ομιλίας. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε MATLAB ή Python (με ισχυρή προτίμηση στη δεύτερη - αν δε γνωρίζετε, θα χρειαστεί να μάθετε τα βασικά).

3. **Ερευνητικό: Ανάλυση και Αναγνώριση Εκφραστικού Περιεχομένου Σημάτων Φωνής με χρήση τεχνικών Βαθιάς Μάθησης - 12 μήνες:** Η αναγνώριση συναισθήματος αποτελεί σημαντικό κεφάλαιο στην Επεξεργασία Φωνής. Υπάρχουν πολλά χαρακτηριστικά (features) της ομιλίας που μπορούν να χρησιμοποιηθούν προς αυτήν την κατεύθυνση αλλά τα τελευταία χρόνια έχουν εισαχθεί μέθοδοι βαθιάς μάθησης που βασίζονται σε νευρωνικά μοντέλα (CNN, LSTM, ResNets, etc) που δε χρησιμοποιούν χαρακτηριστικά. Αντίθετα, βασίζονται σε “ωμές” (raw) μορφές δεδομένων, όπως φασματογραφήματα ή το ίδιο το σήμα της ομιλίας (δείγματα). Θα θέλαμε να χρησιμοποιήσουμε τέτοιες μεθόδους για την αναγνώριση συναισθήματος ομιλητών. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση φωνής και των χαρακτηριστικών της, αρχιτεκτονικών βαθιάς μάθησης, καθώς και γνώση χαρακτηριστικών της εκφραστικής ομιλίας. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε MATLAB ή Python (με ισχυρή προτίμηση στη δεύτερη - αν δε γνωρίζετε, θα χρειαστεί να μάθετε τα βασικά).
4. **Ερευνητικό: Εφαρμογή της Απόκλισης Διαταραχής Φάσης σε Παθολογική Ομιλία - 6 μήνες:** Η Απόκλιση Διαταραχή Φάση (Phase Distortion Deviation - PDD) είναι μια πρόσφατη τεχνική που εφαρμόζεται σε αρμονικά μοντέλα φωνής και μας πληροφορεί για τη διαταραχή της φάσης του έμφωνου - κυρίως - λόγου από την αρμονικότητα και σχετίζεται με το σχήμα του σήματος της γλωττίδας, δηλ. του σήματος που παράγεται ακριβώς πάνω από τις φωνητικές μας χορδές. Η τεχνική αυτή έχει εφαρμοστεί για τη διάκριση ομιλίας με κάποια ιδιαιτερότητα ή παθολογία σε σχέση με κανονικούς ομιλητές μέσω κλασικών τεχνικών στατιστικής ανάλυσης και μηχανικής μάθησης. Θέλουμε να δούμε αν υπάρχει σχέση ανάμεσα στην Απόκλιση Διαταραχής Φάσης και στην ομιλία που προέρχεται από άλλους ομιλητές, αν δηλαδή μπορούμε να διαχωρίσουμε/αναγνωρίσουμε τους ομιλητές με βάση αυτή τη μετρική αλλά με χρήση τεχνικών βαθιάς μάθησης (CNN, LSTM, ResNets, etc), που θα λαμβάνουν διδιάστατες απεικονίσεις της Διαταραχής Φάσης. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση φωνής και ημιτονοειδών μοντέλων, καθώς και χαρακτηριστικών της ομιλίας από παθολογικούς ομιλητές. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε MATLAB.
5. **Ερευνητικό: Εφαρμογή προχωρημένων μεθόδων μηχανικής μάθησης για Αναγνώριση Ομιλίας - 12 μήνες:** Οι τεχνολογίες φωνής όπως αυτή της αυτόματης αναγνώρισης ομιλίας είναι ιδιαίτερα χρήσιμες μεταξύ άλλων στην επικοινωνία ανθρώπου-μηχανής και στην ανάκτηση πληροφορίας σε πολυμεσικά δεδομένα. Τα τελευταία χρόνια, η χρήση προχωρημένων μεθόδων μηχανικής μάθησης έχει βελτιώσει σημαντικά την ακρίβεια των συστημάτων αναγνώρισης ομιλίας. Στην εργασία αυτή θα γίνει χρήση βιβλιοθηκών μηχανικής μάθησης (deep learning) για την αναγνώριση ομιλίας σε βάσεις δεδομένων που έχουν καταγραφεί από συναντήσεις. Θα μάθετε τη διαδικασία δημιουργίας μεγάλου συστήματος διαχείρισης δεδομένων, επικοινωνία των υποσυστημάτων του συστήματος, εφαρμοσμένο προγραμματισμό. Χρειάζεται καλή γνώση σε προγραμματισμό C, Python, script languages. Η εργασία αυτή θα διεκπεραιωθεί σε συνεργασία με το Δρ Βασίλη Τσιάρα, ΕΔΙΠ στο Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Πολυτεχνείο Κρήτης.
6. **Αυτόματη τμηματοποίηση βάσεων δεδομένων φωνής - 6-12 μήνες:** Οι τεχνολογίες φωνής όπως αυτή της αυτόματης αναγνώρισης ομιλίας και σύνθεσης φωνής από κείμενο είναι ιδιαίτερα χρήσιμες μεταξύ άλλων στην επικοινωνία ανθρώπου-μηχανής και στην ανάκτηση πληροφορίας σε πολυμεσικά δεδομένα. Τα τελευταία χρόνια, η χρήση προχωρημένων μεθόδων μηχανικής μάθησης έχει βελτιώσει σημαντικά την ακρίβεια των συστημάτων αναγνώρισης ομιλίας. Για να είναι αποδοτικές αυτές οι μέθοδοι όμως, χρειάζεται να υπάρχουν τμηματοποιημένες σε επίπεδο φωνήματος βάσεις δεδομένων φωνής. Σε αυτή την εργασία θα μάθετε την εφαρμογή επεξεργασίας σήματος για την εξαγωγή παραμέτρων για τη φωνή οι οποίες θα χρησιμοποιηθούν για την αυτόματη τμηματοποίηση με σύγχρονες μεθόδους μηχανικής μάθησης, χρησιμοποιώντας βιβλιοθήκες όπως TensorFlow (Google) κ.λ.π. Χρειάζεται καλή γνώση επεξεργασίας σήματος και προγραμματισμού κυρίως σε Python και σε script languages.

7. **Ανάπτυξη αλγορίθμων αποσύνθεσης σήματος φωνής σε περιδικά και απεριοδικά σήματα - 6/12 μήνες:** Σχετικά πρόσφατα έχουν αναπτυχθεί μέθοδοι αποσύνθεσης ενός σήματος φωνής σε ένα περιοδικό και ένα απεριοδικό σήμα, με χρήση του Διακριτού Αρμονικού Μετασχηματισμού (Discrete Harmonic Transform - DHT). Ένα από τα προσόντα του DHT είναι ότι μπορεί να συγχρονίζει τη συνάρτηση βάσης του με τη χρονικά μεταβαλλόμενη δομή της θεμελιώδους συχνότητας της φωνής. Η παρούσα εργασία στοχεύει στη μελέτη του αλγορίθμου και στην ανάπτυξη ενός συστήματος που αναλύει ένα σήμα φωνής σε μια περιοδική και μια απεριοδική συνιστώσα. Η γνώση που θα αποκτήσετε περιλαμβάνει την ανάλυση και μοντελοποίηση σημάτων φωνής, καθώς και τον προγραμματισμό τους σε MATLAB, Python. Ως εκ τούτου, πολύ καλές γνώσεις επεξεργασίας σήματος και προγραμματισμού σε MATLAB είναι απαραίτητες, ενώ το κατάλληλο υπόβαθρο επεξεργασίας σήματος φωνής θα δοθεί ως πρώτη μελέτη.
8. **Aliasing-free Μοντέλα για το σήμα της γλωττίδας - 6/12 μήνες:** Πολύ πρόσφατα δημοσιεύτηκαν τρόποι να υλοποιεί κανείς μαθηματικά μοντέλα για το σήμα που παράγεται πάνω από τις φωνητικές μας χορδές χωρίς αυτά να έχουν “ψευδώνυμες” συχνότητες κατά τη διακριτοποίησή τους. Οι τρόποι αυτοί εφαρμόστηκαν με επιτυχία στο πιο γνωστό μαθηματικό μοντέλο, το Liljencrants-Fant (LF) model. Χρησιμοποιήθηκε αριθμητική μιγαδικών αριθμών και μετασχηματισμοί Laplace για να εφαρμοστεί η τεχνική στο μοντέλο αυτό. Στο εργαστήριο εφαρμόσαμε τη μέθοδο αυτή και σε άλλα μοντέλα για το σήμα της γλωττίδας, όπως τα Rosenberg, R++, και Klatt model. Στόχος της εργασίας αυτής είναι να υλοποιήσουμε συνθετική ομιλία ώστε να δούμε αν τα μοντέλα αυτά έχουν και αντιληπτική βελτίωση, πέρα από αλγοριθμική. Η γνώση που θα αποκτήσετε περιλαμβάνει την ανάλυση και μαθηματική μοντελοποίηση σημάτων φωνής, καθώς και τον προγραμματισμό τους σε MATLAB. Πολύ καλές γνώσεις επεξεργασίας σήματος και προγραμματισμού σε MATLAB είναι απαραίτητες, ενώ το κατάλληλο υπόβαθρο επεξεργασίας σήματος φωνής θα δοθεί ως πρώτη μελέτη.
9. **Ανάπτυξη αλγορίθμου εξαγωγής της φασματικής περιβάλλουσας του σήματος φωνής παρουσία θορύβου - 6/12 μήνες:** Πολύ πρόσφατα, ερευνητές ανέπτυξαν μια μέθοδο εξαγωγής της φασματικής περιβάλλουσας ενός σήματος φωνής η οποία είναι εύρωστη σε θόρυβο και βασίζεται στην κλασική τεχνική της Γραμμικής Πρόβλεψης (Linear Prediction). Στην παρούσα εργασία, ζητείται η υλοποίηση της μεθόδου time-regularized linear prediction η οποία έχει αποδειχθεί ότι είναι ιδιαίτερα αποδοτική. Η γνώση που θα αποκτήσετε περιλαμβάνει την ανάλυση και μοντελοποίηση σημάτων φωνής, τη μοντελοποίηση της παραγωγής ανθρώπινης ομιλίας και τα στάδιά της, τα σήματα που εμπλέκονται στην παραγωγή αυτή, καθώς και τον προγραμματισμό τους σε MATLAB. Πολύ καλές γνώσεις επεξεργασίας σήματος και προγραμματισμού σε MATLAB είναι απαραίτητες, ενώ το κατάλληλο υπόβαθρο επεξεργασίας σήματος φωνής θα δοθεί ως πρώτη μελέτη.
10. **Ερευνητικό: Πρόβλεψη Προσωπικότητας Ομιλητή με μεθόδους μηχανικής μάθησης - 12 μήνες:** Από πλευράς ψυχολογίας, η προσωπικότητα ενός ανθρώπου είναι συνάρτηση πέντε χαρακτηριστικών: *ευσυνειδησία, ειλικρίνεια, εξωστρέφεια, τερπνότητα, και νευρωτισμός*. Η πρόβλεψη προσωπικότητας ενός ομιλητή από ένα τμήμα ομιλίας του είναι ιδιαίτερα χρήσιμη σε καταστάσεις που θέλουμε να προβλέψουμε τη μελλοντική συμπεριφορά του ή αν ταιριάζει σε μια θέση που θέλουμε να του εμπιστευτούμε (η οποία θεωρούμε ότι πρέπει να έχει συγκεκριμένο βάρος από κάθε χαρακτηριστικό), ακόμα και σε εφαρμογές dating ή monitoring της ψυχικής κατάστασης ενός ομιλητή. Η πρόβλεψη αυτή γίνεται εξάγοντας χαρακτηριστικά της ομιλίας και εφαρμόζοντας στατιστικά επάνω σε αυτά. Στη βιβλιογραφία συνηθίζεται να ταξινομούνται οι ομιλητές με βάση την δυαδική ένταση (high/low) καθενός από τα πέντε χαρακτηριστικά. Όμως οι αξιολογήσεις γίνονται με βάση scores σε μια συνεχή κλίμακα 1-5. Σε αυτήν την πτυχιακή θα θέλαμε να δούμε μοντέλα παλινδρόμησης (regression) και όχι ταξινόμησης (classification), με στόχο την πρόβλεψη του score του κάθε ομιλητή σε κάθε χαρακτηριστικό. Θα χρησιμοποιηθεί η βάση δεδομένων SSPNet αλλά θα εξεταστεί η πιθανότητα να εμπλουτιστεί με παραπάνω (Ελληνες) αξιολογητές.

Η γνώση που θα αποκτήσετε περιλαμβάνει τεχνικές μηχανικής μάθησης και προχωρημένο προγραμματισμό σε Python. Πολύ καλές γνώσεις επεξεργασίας σήματος και προγραμματισμού είναι απαραίτητες, ενώ το κατάλληλο υπόβαθρο ψυχολογίας θα δοθεί ως πρώτη μελέτη.

11. **Ερευνητικό: Βελτιστοποίηση Γραφικού Περιβάλλοντος με Προσθήκη Εξαγωγής Χαρακτηριστικών για Πρόβλεψη Προσωπικότητας Ομιλητή - 12 μήνες:** Από πλευράς ψυχολογίας, η προσωπικότητα ενός ανθρώπου είναι συνάρτηση πέντε χαρακτηριστικών: *ευσυνειδησία, ειλικρίνεια, εξωστρέφεια, τερπνότητα, και νευρωτισμός*. Η πρόβλεψη προσωπικότητας ενός ομιλητή από ένα τμήμα ομιλίας του είναι ιδιαίτερα χρήσιμη σε καταστάσεις που θέλουμε να προβλέψουμε τη μελλοντική συμπεριφορά του ή αν ταιριάζει σε μια θέση που θέλουμε να του εμπιστευτούμε (η οποία θεωρούμε ότι πρέπει να έχει συγκεκριμένο βάρος από κάθε χαρακτηριστικό), ακόμα και σε εφαρμογές dating ή monitoring της ψυχικής κατάστασης ενός ομιλητή. Η πρόβλεψη αυτή γίνεται εξάγοντας χαρακτηριστικά της ομιλίας και εφαρμόζοντας στατιστικά επάνω σε αυτά. Στη βιβλιογραφία συνηθίζεται να ταξινομούνται οι ομιλητές με βάση την δυαδική ένταση (high/low) καθενός από τα πέντε χαρακτηριστικά. Σε αυτήν την πτυχιακή θέλουμε να χρησιμοποιήσουμε έναν ήδη υπάρχον γραφικό περιβάλλον και να του προσθέσουμε κάποιες επιπλέον λειτουργίες. Για παράδειγμα, ο πρώτος στόχος θα είναι να μπορεί να εξάγει χαρακτηριστικά από ομιλητές (αυτή τη στιγμή παίρνει έτοιμα κάποια χαρακτηριστικά). Θα χρησιμοποιηθεί η βάση δεδομένων SSPNet αλλά θα εξεταστεί η πιθανότητα να εμπλουτιστεί με παραπάνω (Έλληνες) αξιολογητές. Η γνώση που θα αποκτήσετε περιλαμβάνει τεχνικές μηχανικής μάθησης και προχωρημένο προγραμματισμό σε Python. Πολύ καλές γνώσεις επεξεργασίας σήματος και προγραμματισμού είναι απαραίτητες, ενώ το κατάλληλο υπόβαθρο ψυχολογίας θα δοθεί ως πρώτη μελέτη.
12. **Ερευνητικό: Αναγνώριση συναισθήματος ομιλητή σε περιβάλλοντα θορύβου με τεχνικές βαθιάς μάθησης - 12 μήνες:** Η αναγνώριση συναισθήματος αποτελεί σημαντικό κεφάλαιο στην Επεξεργασία Φωνής. Υπάρχουν πολλά χαρακτηριστικά (features) της ομιλίας που μπορούν να χρησιμοποιηθούν προς αυτήν την κατεύθυνση. Όμως στην πραγματικότητα μπορεί να υπάρχει (ισχυρός ή μη) θόρυβος στο περιβάλλον που επιδιώκουμε την αναγνώριση του συναισθήματος, ο οποίος αλλοιώνει τα χαρακτηριστικά αυτά κάνοντάς τα λιγότερο χρήσιμα. Θα θέλαμε να χρησιμοποιήσουμε πιο εύρωστες τεχνικές βαθιάς μάθησης αντί των χαρακτηριστικών αυτών (CNN, LSTM, etc) για την αναγνώριση συναισθήματος ομιλητών. Οι τεχνικές θα εφαρμοστούν σε πλειάδα βάσεων δεδομένων για καλύτερη αξιολόγηση. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση φωνής και των χαρακτηριστικών της, τεχνικών βαθιάς μάθησης, καθώς και γνώση χαρακτηριστικών της εκφραστικής ομιλίας. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε MATLAB ή Python (με ισχυρή προτίμηση στη δεύτερη - αν δε γνωρίζετε, θα χρειαστεί να μάθετε τα βασικά).
13. **Ερευνητικό: Αναγνώριση του Covid19 με χρήση χαρακτηριστικών ομιλίας και τεχνικών μηχανικής ή βαθιάς μάθησης - 12 μήνες:** Στην εποχή μας υπάρχει ιδιαίτερη ανάγκη για pre-screening ασθενών που είναι φορείς του ιού SARS-Cov2-19 που μεταδίδει την ασθένεια Covid19. Τόσο η ομιλία όσο και άλλα χαρακτηριστικά (αναπνοή, βήχας, κλπ) αποτελούν στοιχεία που μπορούν να χρησιμοποιηθούν για την πρόβλεψη της ασθένειας. Θα θέλαμε να δούμε τόσο κλασικές τεχνικές μηχανικής μάθησης επάνω σε χαρακτηριστικά που εξάγονται από την ομιλία όσο και αρχιτεκτονικές βαθιάς μάθησης επάνω σε περισσότερο “ωμές” εκφράσεις των δεδομένων, με σκοπό την πρόβλεψη της ασθένειας από στοιχεία της ομιλίας. Υπάρχει πρόσφατη βιβλιογραφία την οποία θα μελετήσετε και θα προσπαθήσετε να εξελίξετε. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση φωνής και των χαρακτηριστικών της, τεχνικών μηχανικής ή βαθιάς μάθησης, καθώς και γνώση χαρακτηριστικών της ομιλίας των ασθενών με Covid19. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε MATLAB ή Python.

14. **Ερευνητικό: Ταξινόμηση Παθολογιών με χρήση Ακουστικής Ανάλυσης Αναπνευστικών Σημάτων και Τεχνικών Μηχανικής ή Βαθιάς Μάθησης - 12 μήνες:** Στην εποχή μας υπάρχει ιδιαίτερη ανάγκη για pre-screening ασθενών που είναι φορείς ασθνευιών όπως η φυματίωση (tuberculosis). Τόσο η ομιλία όσο και άλλα χαρακτηριστικά (αναπνοή, βήχας, κλπ) αποτελούν στοιχεία που μπορούν να χρησιμοποιηθούν για την πρόβλεψη της ασθένειας. Θα θέλαμε να δούμε τόσο κλασικές τεχνικές μηχανικής μάθησης επάνω σε χαρακτηριστικά που εξάγονται από την ομιλία ή το βήχα, όσο και αρχιτεκτονικές βαθιάς μάθησης (WideResNets, DiracNets) επάνω σε περισσότερο “ωμές” εκφράσεις των δεδομένων, με σκοπό την πρόβλεψη της ασθένειας από στοιχεία της ομιλίας. Υπάρχει πρόσφατη βιβλιογραφία την οποία θα μελετήσετε και θα προσπαθήσετε να εξελίξετε. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση φωνής και των χαρακτηριστικών της, τεχνικών μηχανικής ή βαθιάς μάθησης. Απαιτείται καλή γνώση επεξεργασίας σήματος, μηχανικής μάθησης, και προγραμματισμού σε Python.
15. **Ερευνητικό: Χρήση Φασμάτων Διαμόρφωσης (modulation spectra) στην Ταξινόμηση Ασθνευιών με χρήση Ακουστικής Ανάλυσης Αναπνευστικών Σημάτων και Τεχνικών Μηχανικής ή Βαθιάς Μάθησης - 12 μήνες:** Στην εποχή μας υπάρχει ιδιαίτερη ανάγκη για pre-screening ασθενών που είναι φορείς ασθνευιών όπως η φυματίωση (tuberculosis). Τόσο η ομιλία όσο και άλλα χαρακτηριστικά (αναπνοή, βήχας, κλπ) αποτελούν στοιχεία που μπορούν να χρησιμοποιηθούν για την πρόβλεψη της ασθένειας. Στόχος της εργασίας αυτής είναι να εφαρμοστεί η τεχνική των Φασμάτων Διαμόρφωσης, που ουσιαστικά αποτελούνται από δυο διαδοχικούς short-time Fourier Transforms, για την αναπαράσταση των σημάτων. Στο παρελθόν αυτή η τεχνική έχει αποδειχθεί χρήσιμη στην ταξινόμηση παθολογιών. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση σήματος και των χαρακτηριστικών της με χρήση φασμάτων διαμόρφωσης, τεχνικών μηχανικής ή βαθιάς μάθησης. Απαιτείται καλή γνώση επεξεργασίας σήματος, μηχανικής μάθησης, και προγραμματισμού σε Python.
16. **Περίληψη και Σύγκριση Αλγορίθμων Ανίχνευσης Έναρξης Ακουστικών Γεγονότων - 12 μήνες:** Στη βιβλιογραφία υπάρχουν πολλοί αλγόριθμοι εύρεσης έναρξης ακουστικών γεγονότων, όπως το χτύπημα δυο πιάτων, ένα φτάρνισμα, ένα γαύγισμα, ένα σφύριγμα, κλπ. Οι αλγόριθμοι αυτοί βασίζονται τόσο σε χρονικά όσο και σε φασματικά χαρακτηριστικά ηχητικών σημάτων. Στόχος της πτυχιακής είναι η βιβλιογραφική έρευνα και η υλοποίηση τέτοιων αλγορίθμων, καθώς και η σύγκρισή τους σε επίπεδο πολυπλοκότητας, ταχύτητας, μνήμης, και απόδοσης. Η γνώση που θα αποκτήσετε περιλαμβάνει ανάλυση ηχητικών σημάτων και των χαρακτηριστικών της και τεχνικών επεξεργασίας σήματος. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε Python.
17. **Ανάπτυξη Διεπαφής για Παρουσίαση Χρονικών και Φασματικών Μετρικών από Ακουστική Ανάλυση Σημάτων - 12 μήνες:** Στη Φωνητική και στη Γλωσσολογία, συνηθίζεται να αναλύεται η γλώσσα και τα φωνήματά της μέσω ηχογραφήσεων. Πολλές μετρικές έχουν προταθεί για την ανάλυση φωνημάτων με σκοπό τη σύγκριση ομιλητών, διαλέκτων, παθολογιών, και γλωσσών. Μια τέτοια ανάλυση απαιτεί εξειδικευμένες γνώσεις και εργαλεία όπως το MATLAB. Στόχος της εργασίας είναι η δημιουργία μιας διεπαφής που θα είναι απλή και εύχρηστη και θα απευθύνεται στο μη ειδικό χρήστη (όπως ένας γλωσσολόγος), και θα παρουσιάζει κάποιες μετρικές με βάση ηχητικά σήματα που δίνονται ως είσοδο στη διεπαφή. Σκοπός είναι η δημιουργία ενός εργαλείου αυτόματης ανάλυσης τμημάτων της γλώσσας όπως τα φωνήεντα και τα σύμφωνα. Απαιτείται καλή γνώση επεξεργασίας σήματος και προγραμματισμού σε Python.