

# Βιοϊατρική Τεχνολογία

## Εργασία-Project 2018-2019

Στην εργασία αυτή θα ασχοληθείτε με τον εντοπισμό και την ανάλυση κορυφών (spikes) από εξωκυτταρικές καταγραφές του εγκεφάλου.

Στο φάκελο της εργασίας που σας μοιράστηκε θα βρείτε ένα υπο-φάκελο με την ονομασία 'Data' και ένα υπο-φάκελο με την ονομασία 'Code'. Στους υπο-φάκελους αυτούς υπάρχουν αρχεία δεδομένων στη μορφή .mat και αρχείο κώδικα σε Matlab, αντίστοιχα, τα οποία θα χρησιμοποιήσετε στην εργασία όπως περιγράφεται παρακάτω.

### Θέμα 1<sup>ο</sup>

Για την ανίχνευση κορυφών από εξωκυτταρικές καταγραφές που έχουν υποστεί ζωνοδιαβατό φιλτράρισμα 300-3000Hz, χρησιμοποιείται συνήθως η μέθοδος της κατωφλίωσης. Πιο συγκεκριμένα χρησιμοποιώντας ένα κατώφλι πλάτους, εντοπίζεται και καταγράφεται ως κορυφή η δραστηριότητα του σήματος όπου σε μια γειτονιά υπερβαίνει την τιμή ενός κατωφλίου  $T$ .

Τα αρχεία Data\_Test\_ $x$  όπου  $x = 1, \dots, 8$ , περιέχουν διαφορετικά σήματα εξωκυτταρικής καταγραφής που έχουν υποστεί ζωνοδιαβατό φιλτράρισμα. Σε κάθε αρχείο Data\_Test\_ $x$  εμπεριέχεται τόσο το σήμα (διάνυσμα με ονομασία 'data') όσο και μια μεταβλητή με την ονομασία spikeNum η οποία υποδηλώνει τον πραγματικό αριθμό των κορυφών που βρίσκεται σε κάθε σήμα 'data'.

Ερώτημα 1.1: Απεικονίστε τα 10000 πρώτα δείγματα των σημάτων που βρίσκονται τα αρχεία Data\_Test\_ $x$  (ένα γράφημα για κάθε αρχείο). Τι παρατηρείτε;

Μια σχέση που χρησιμοποιείται συνήθως στη βιβλιογραφία για τον ορισμό της τιμής του κατωφλίου για τον εντοπισμό κορυφών είναι η σχέση  $T = k \cdot \sigma_n$  όπου  $\sigma_n$  είναι μια εκτίμηση της τυπικής απόκλισης του θορύβου (πληροφορία άσχετη με τη δραστηριότητα των κορυφών).

Ερώτημα 1.2: Χρησιμοποιώντας τη σχέση  $T = k \cdot \sigma_n$  για διάφορες τιμές της παραμέτρου  $k$  καταμετρήστε με τη μέθοδο της κατωφλίωσης τις κορυφές που υπάρχουν σε κάθε ένα από τα σήματα στα αρχεία Data\_Test\_ $x$ . Κάντε το γράφημα **#κορυφών vs.  $k$**  και για τα οκτώ σήματα (ένα γράφημα για κάθε αρχείο). Τι παρατηρείτε;

Ερώτημα 1.3: Ορίσετε έναν εμπειρικό κανόνα για την παράμετρο  $k$  ώστε ο αριθμός των κορυφών που υπολογίσατε για κάθε σήμα εξωκυτταρικής καταγραφής να είναι κοντά στον πραγματικό αριθμό (μεταβλητή spikeNum). Ποιος είναι ο κανόνας αυτός;

## Θέμα 2°

Τα αρχεία Data\_Eval\_E\_x όπου  $x = 1, \dots, 4$  περιέχουν διαφορετικά σήματα εξωκυτταρικής καταγραφής. Το σήμα της καταγραφής υπάρχει πάλι στο διάνυσμα με το όνομα 'data'.

Ερώτημα 2.1: Χρησιμοποιώντας τον εμπειρικό κανόνα του ερωτήματος 1.3 εντοπίστε τις κορυφές σε κάθε ένα από τα σήματα στα αρχεία Data\_Eval\_E\_x. Πόσες κορυφές εντοπίζετε σε κάθε ένα από τα σήματα των αρχείων Data\_Eval\_E\_x; Καταχωρίστε τις χρονικές στιγμές που τις εντοπίζετε σε ένα διάνυσμα με όνομα spikeTimesEst (ένα διάνυσμα για κάθε αρχείο).

Ερώτημα 2.2: Απομονώστε τις κυματομορφές των κορυφών που εντοπίσατε σε κάθε αρχείο και στοιχίστε τις με βάση το πρώτο ακρότατό τους σε ένα πίνακα SpikesEst (ένα πίνακα για κάθε αρχείο); Απεικονίστε όλες τις κυματομορφές των κορυφών που εντοπίσατε σε ένα γράφημα (ανά αρχείο δεδομένων. Φροντίστε κάθε κυματομορφή να απεικονίζεται πλήρως). Τι παρατηρείτε για  $x = 1, \dots, 4$ ;

Στα αρχεία Data\_Eval\_E\_x εκτός από τα σήματα (διάνυσμα 'data') εμπεριέχεται και το διάνυσμα spikeTimes. Η μεταβλητή spikeTimes δείχνει τις χρονικές στιγμές που εμφανίζονται οι κορυφές (spikes) σε κάθε ένα από τα σήματα των αρχείων Data\_Eval\_E\_x.

Ερώτημα 2.3: Με βάση τη μεταβλητή spikeTimes υπολογίστε πόσες κορυφές εντοπίσατε σωστά και πόσες από τις κορυφές που εντοπίσατε δεν είναι στην πραγματικότητα κορυφές αλλά δραστηριότητα που δεν σχετίζεται με τη δραστηριότητα νευρώνων (θόρυβος);

*Παρατήρηση: Ενδέχεται οι χρονικές στιγμές που καταχωρίσατε στο spikeTimesEst του ερωτήματος 2.1 να διαφέρουν ελάχιστα από αυτές που αναγράφονται στο spikeTimes παρόλο που αναφέρονται στις ίδιες κορυφές. Προσαρμόστε ανάλογα τον κώδικά σας για να κάνετε την αντιστοίχιση μεταξύ των πραγματικών κορυφών και των κορυφών που εντοπίσατε εσείς.*

Τα σήματα που βρίσκονται στα αρχεία Data\_Eval\_E\_x περιέχουν την δραστηριότητα τριών νευρώνων. Δηλαδή, οι κορυφές που εντοπίσατε αντιστοιχούν σε τρεις διαφορετικούς νευρώνες. Στο διάνυσμα spikeClass των αρχείων Data\_Eval\_E\_x θα βρείτε για κάθε μία από τις πραγματικές κορυφές (spikeTimes) το νούμερο του νευρώνα στον οποίο αντιστοιχεί, 1, 2, ή 3.

Ερώτημα 2.4: Για κάθε μία από τις κυματομορφές που εντοπίσατε και καταχωρίσατε στον πίνακα SpikesEst για κάθε αρχείο Data\_Eval\_E\_x, υπολογίστε δύο χαρακτηριστικά της (π.χ. μέγιστο πλάτος κυματομορφής και αριθμός διαβάσεων από το μηδέν). Επιλέξτε δύο χαρακτηριστικά που εσείς επιθυμείτε (δεν υπάρχει περιορισμός). Απεικονίστε σαν σημεία στον δισδιάστατο χώρο τα ζεύγη των χαρακτηριστικών που υπολογίσατε για κάθε κυματομορφή χρησιμοποιώντας διαφορετικό χρώμα για κάθε νευρώνα (ένα γράφημα για κάθε αρχείο). Τι παρατηρείτε για τα διαφορετικά  $x = 1, \dots, 4$ ;

Στον υπο-φάκελο 'Code' υπάρχει μια ρουτίνα MyClassify(Data, group) σε κώδικα Matlab η οποία υλοποιεί μια γραμμική μεθοδολογία μηχανικής μάθησης. Παίρνει ως είσοδο έναν πίνακα Data και ένα διάνυσμα group. Κάθε γραμμή του πίνακα Data πρέπει να περιέχει τα χαρακτηριστικά που υπολογίσατε για κάθε κυματομορφή στο ερώτημα 2.4. Το διάνυσμα group πρέπει να περιέχει το νούμερο κάθε νευρώνα στον οποίο αντιστοιχεί η κυματομορφή από την οποία υπολογίσατε τα χαρακτηριστικά του πίνακα Data. Δηλαδή, αν η πρώτη γραμμή του πίνακα Data περιέχει χαρακτηριστικά που υπολογίστηκαν από κυματομορφή του νευρώνα 2, το πρώτο στοιχείο του διανύσματος group πρέπει να είναι ο αριθμός 2. Αν η δεύτερη γραμμή του πίνακα Data αντιστοιχεί σε κυματομορφή του νευρώνα 1 το δεύτερο στοιχείο του διανύσματος group πρέπει να είναι ο αριθμός 1 κοκ.

Η συνάρτηση MyClassify επιστρέφει το ποσοστό ταξινόμησης που επιτυγχάνεται χρησιμοποιώντας τα χαρακτηριστικά που εισαγάγατε με βάση ένα γραμμικό αλγόριθμο μηχανικής μάθησης.

Ερώτημα 2.5: Δημιουργήστε τον πίνακα Data και το διάνυσμα group για κάθε ένα από τα αρχεία Data\_Eval\_E\_x. Έπειτα, δώστε τα σαν είσοδο στην συνάρτηση MyClassify και πάρτε την έξοδο που αφορά στο ποσοστό σωστής ταξινόμησης (αναγνώρισης) των τριών νευρώνων. Τι ποσοστό επιτυγχάνετε για τα διαφορετικά αρχεία  $x = 1, \dots, 4$ ; Αυξήστε τον αριθμό των χαρακτηριστικών και επαναλάβετε το πείραμα (και πάλι επιλέξτε όποια χαρακτηριστικά εσείς επιθυμείτε – μέγιστο δέκα χαρακτηριστικά). Υπάρχει βελτίωση στα ποσοστά σωστής αναγνώρισης των νευρώνων;

## Βιβλιογραφική Εργασία 2018-2019

Στη βιβλιογραφική εργασία θα ασχοληθείτε με ένα θέμα της επιλογής σας το οποίο θα αναπτύξετε σε τέσσερεις βασικούς τομείς (ενδεικτική δομή εργασίας):

- Περιγραφή προβλήματος/θέματος
- Σημασία του προβλήματος/θέματος
- Ενδεικτικές μεθοδολογίες και περιγραφή
- Ενδεικτικές εφαρμογές και περιγραφή

Η δομή που περιεγράφηκε παραπάνω είναι ενδεικτική και μπορείτε να την προσαρμόσετε αναλόγως στο θέμα που θα επιλέξετε.

Η βιβλιογραφική αναφορά που θα καταθέσετε θα είναι 15-30 σελίδες συμπεριλαμβανομένων το αναφορών.

Παρακάτω δίνονται ενδεικτικά θέματα για τη βιβλιογραφική εργασία:

Θέμα
Εντοπισμός επιληψίας σε ηλεκτροεγκεφαλογράφημα
Συστήματα P300 spellers (με EEG)
Συστήματα αλληλεπίδρασης εγκεφάλου υπολογιστή με σήματα fNIRS
Αναγνώριση αρρυθμίας από ηλεκτροκαρδιογράφημα
Μεθοδολογίες ανάλυσης Αξονικής Τομογραφίας για εντοπισμό όγκων

**Η υλοποίηση των εργασιών θα γίνει σε ομάδες των τριών ατόμων (μέγιστο). Πρέπει να μπείτε στην πλατφόρμα e-learning και να δηλώσετε την ομάδα σας.**

**Παράδοση και των δύο εργασιών σε ένα αρχείο pdf. Στο τέλος του αρχείου να υπάρχει παράρτημα με τον κώδικα σε Matlab που αναπτύξατε για την εργασία-project**

**Παράδοση έως 31-12-2018, 23:59:59 μέσω του συστήματος e-learning**