



26/7/2016

ΠΡΟΗΓΜΕΝΕΣ ΜΕΘΟΔΟΙ ΕΠΕΞΕΡΓΑΣΙΑΣ ΣΗΜΑΤΟΣ 8ο ΕΞΑΜΗΝΟ

I. Εισαγωγικά

Σκοπός Πειραμάτων

Ο κοινός σκοπός των πειραμάτων που διεξήχθησαν ήταν η καταγραφή της ηλεκτροεγκεφαλικής δραστηριότητας (ηλεκτροεγκεφαλογράφημα (ΗΕΓ)), του Heart Rate, της αντίστασης του δέρματος (R) και της θερμοκρασίας του δέρματος (T) των συμμετεχόντων κατά τη διάρκεια θέασης βίντεο που σχετίζονταν με το bullying:

Π1. Θέαση βίντεο σε laptop

Π2. Θέαση βίντεο σε Virtual Reality (VR) με βίντεο 360°.

Συσκευές Καταγραφής ΗΕΓ, ΗΒ, R, T

Σε κάθε ένα από τα πειράματα, η καταγραφή της ηλεκτροεγκεφαλικής δραστηριότητας πραγματοποιήθηκε με χρήση της συσκευής Clinical Geodesic EEG System 400 (<http://www.egi.com/clinical-division/clinical-division-clinical-products/ges-400-series>), η οποία επιτρέπει την ταυτόχρονη καταγραφή 256 καναλιών ΗΕΓ με συχνότητα δειγματοληψίας **250 Hz**.

Επίσης, οι καταγραφές του HRV, R and T έγιναν από το MS Band 2 (<https://www.microsoft.com/microsoft-band/en-us>) με συχνότητα δειγματοληψίας **8 Hz**.

Πειραματική Διαδικασία

Δεδομένα

Τα πειραματικά δεδομένα αντιστοιχούν σε 18 συμμετέχοντες (Αγόρια: #5-#10, #15, #17-#20, Κορίτσια: #2-#4, #11-#13, #16) και αριθμούνται με το πρόθεμα ptes#.mat. Για κάθε συμμετέχοντα υπάρχουν καταγραφές από κατάσταση ηρεμίας (φάκελος matfilesHV), καταγραφές από το πείραμα Π1 (φάκελος "matfilesT1") και καταγραφές από το πείραμα Π2 (φάκελος "matfilesT2"). Κάθε αρχείο καταγραφής των πειραμάτων Π1 και Π2 περιλαμβάνει αρχεία που αφορούν στην απόκριση ΗΕΓ κατά τη θέαση στιγμών bullying που ξεκινούν από bul, καθώς και αρχεία που αφορούν στην απόκριση ΗΕΓ κατά τη θέαση στιγμών του βίντεο που δεν περιέχουν bullying και τα οποία ξεκινούν από nobul.

Τα segmented αρχεία έχουν όλα διάσταση 257x375, όπου οι πρώτες 1:256 γραμμές αντιστοιχούν στα 256 κανάλια, ενώ η γραμμή 257 αντιστοιχεί στο κανάλι αναφοράς (μπορείτε να το αγνοήσετε). Όλα τα segments έχουν το ίδιο μήκος (800 ms).



Σε κάθε συμμετέχοντα αντιστοιχεί επίσης ένα αρχείο της μορφής ptes#events στο φάκελο “events”. Στα αρχεία αυτά, οι κωδικοί HV, TS1 και TS2 αντιστοιχούν στην κατάσταση ηρεμίας, στο πείραμα Π1 και στο πείραμα Π2, αντίστοιχα, ενώ οι κωδικοί D254 και D255 αντιστοιχούν σε στιγμές bullying και σε στιγμές που δεν περιέχουν bullying, αντίστοιχα. Η προτελευταία στήλη των αρχείων αυτών, με τίτλο onset, περιέχει τη στιγμή έναρξης του αντίστοιχου σταδίου του πειράματος/ερεθίσματος, στη μορφή ώρα:λεπτά:δευτερόλεπτα:χιλιοστά.

Επιπλέον, στο φάκελο “matfiles_band” σας δίνεται για κάθε συμμετέχοντα ένα αρχείο (ptes#band) που περιέχει τις μετρήσεις του HeartRate, της αντίστασης του δέρματος και τις θερμοκρασίες του δέρματος κατά τη διάρκεια των καταγραφών, μαζί με τις αντίστοιχες χρονικές τους στιγμές. Οι στήλες των .mat αρχείων αντιστοιχούν σε Χρονιά, Μήνα, Ημέρα, Ώρα, Λεπτό, Δευτερόλεπτο, HeartRate, Αντίσταση Δέρματος και Θερμοκρασία Δέρματος, όπως δηλώνεται και στο αρχείο BandFieldsInfo.mat

Τέλος, σας δίνεται αρχείο αντιστοίχισης των καναλιών με την τοπολογία του εγκεφάλου και την τυποποίηση του συστήματος 10/20 (EEGChannelMapping.rtf), καθώς και μια εικόνα με την τοπολογία των 256 καναλιών.

Προεπεξεργασία

Η προεργασία που έγινε είναι φιλτράρισμα στα 0.3-30 Hz, artifact detection, bad channel replacement, baseline correction και segmentation.

Ανάλυση Σημάτων

Στην ανάλυση των δεδομένων δίνονται **πέντε κατευθύνσεις που αντιστοιχούν σε πέντε ομάδες** που θα δημιουργηθούν από τους φοιτητές. Η οργάνωση και η επιλογή των κατευθύνσεων αφήνονται στη συνεννόηση των φοιτητών μεταξύ τους (π.χ. μέσω κοινού doodle). Για κάθε κατεύθυνση απαιτείται βιβλιογραφική αναδρομή σχετικά με:

1. Περιοχές του εγκεφάλου που ενεργοποιούνται κατά τη διάρκεια των συναισθηματικών διεγέρσεων (π.χ., <http://stanford.edu/~knutson/ans/ledoux00.pdf>)
2. Συχνотικό περιεχόμενο που μας ενδιαφέρει, σύμφωνα με την κατηγοριοποίηση των ζωνών δραστηριότητας του εγκεφάλου (π.χ., <https://en.wikipedia.org/wiki/Electroencephalography>, <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3198250/>)
3. Evoked Related Potentials (ERPs) (π.χ., https://en.wikipedia.org/wiki/Event-related_potential, <http://www.ncbi.nlm.nih.gov/pubmed/22607397>) και ειδικότερα τα early posterior positivity (EPN), N170, P300 και late positive potential (LPP).

Κατεύθυνση 1.-Ανάλυση ERPs #1

- a. Δημιουργείτε ένα σήμα μέσου όρου κάνοντας averaging στα σήματα ανά πείραμα (Π1 και Π2) ανά κατάσταση (bul και nobul) και ανά κανάλι, παράγοντας τα αντίστοιχα ERPs για κάθε συμμετέχοντα. Επικεντρωθείτε στις χρονικές περιοχές



γύρω από το early posterior positivity (EPN), N170, P300 και late positive potential (LPP) χρησιμοποιώντας κατάλληλο χρονικό παράθυρο. Επιπλέον, παραθυροποιείτε και τα σήματα του smartwatch στα ίδια χρονικά διαστήματα με αυτά του EEG. Στα σήματα που προκύπτουν να υπολογιστεί το πλάτος του Bispectrum (μόνο στην πρωτεύουσα περιοχή). Αφήνεται ελεύθερη η επιλογή των παραμέτρων υπολογισμού του Bispectrum με στόχο την καλύτερη εκτίμησή του. Εάν χρειαστεί κάντε κατάλληλο interpolation (εντολή interp1.m) για να έχετε περισσότερα δείγματα στον υπολογισμό του bispectrum.

b. Για κάθε Bispectrum που υπολογίσατε βρείτε

- (i) την τιμή του μεγίστου του Bispectrum
- (ii) την τιμή των (f1,f2) που αντιστοιχούν στο μέγιστο του Bispectrum
- (iii) το std των συχνοτήτων γύρω από την κορυφή του μεγίστου του Bispectrum (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από το μέγιστο που ξεπερνούν το 10% της τιμής του μεγίστου)
- (iv) την τιμή της επόμενης κορυφής (αμέσως μικρότερη της μέγιστης) του Bispectrum
- (v) την τιμή των (f1,f2) που αντιστοιχούν στην τιμή της επόμενης κορυφής του Bispectrum
- (vi) το std των συχνοτήτων γύρω από την επόμενη κορυφή του μεγίστου του Bispectrum (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από την επόμενη κορυφή που ξεπερνούν το 10% της τιμής της επόμενης κορυφής).

c. Φτιάξτε και το κατάλληλο διάνυσμα χαρακτηριστικών που να αποτελείται από τις τιμές που βρήκατε στο b (δηλ. (i)-(vi)), για κάθε ένα συμμετέχοντα ανά κανάλι και ανά πείραμα και ανά κλάση (Π1-bul/nobul, Π2-bul/nobul) για να τα χρησιμοποιείτε ως δεδομένα στον ταξινομητή του επόμενου ερωτήματος.

d. Εφαρμόζοντας Quadratic Discriminant Analysis (QDA) χρησιμοποιώντας την εντολή του Matlab classify.m (<http://www.mathworks.com/help/toolbox/stats/classify.html>) κρατήστε το 75% των δεδομένων για training και το 25% των δεδομένων για testing με τυχαίο τρόπο 10 φορές και βγάλτε τη μέση ακρίβεια στην εκτίμηση των δύο κλάσεων (bul/nobul) ανά πείραμα (Π1, Π2) **ανεξάρτητα του συμμετέχοντα (user independent case)**. Σχολιάστε την εκτίμησή σας για το ποια κατηγορία χαρακτηριστικών παρέχει τη μέγιστη δυνατότητα διαφοροποίησης μεταξύ των δύο κλάσεων.

e. Επαναλάβετε τα βήματα c-d, λαμβάνοντας υπόψη την επίδραση του καναλιού, και με βάση τη βιβλιογραφία, επιλέξτε τα κανάλια που μπορεί να σας αυξήσουν την απόδοση της ταξινόμησης. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

f. Επαναλάβετε τα βήματα c-e, λαμβάνοντας υπόψη την επίδραση του φύλου στα δεδομένα σας. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

g. Συγκρίνετε τα αποτελέσματά σας ανά επιλεγμένο ERP και περιγράψτε τυχόν συσχετίσεις, αλληλεπιδράσεις, αλληλουχίες (π.χ., εάν επηρεάζει κάποιο ERP το επόμενο λόγω της χρονικής αλληλουχίας τους).



h. Εξετάστε εάν υπάρχουν στατιστικώς σημαντικές διαφορές στα διανύσματα χαρακτηριστικών του c. κατά μήκος των πειραμάτων Π1 και Π2 ανά κατηγορία bul και nobul ενοποιώντας τα δεδομένα των χρηστών για το καλύτερο (ή καλύτερα) κανάλι (κανάλια) όπως προκύπτει από το e., για το φύλλο που σας δίνει την καλύτερη ταξινόμηση από το f. και για το ERP που σας δίνει τα καλύτερα αποτελέσματα διαχωρισμού (με βάση το g), ώστε να φανεί η τυχόν επίδραση της εικονικής πραγματικότητας και του βίντεο 360 στη βιωματική αντίληψη του φαινομένου του bullying. Σε δεύτερη προσέγγιση, διαχωρίστε τα δεδομένα από το EEG με αυτά από το smartwatch και δείτε εάν αυτόνομα παρουσιάζουν στατιστικώς σημαντική διαφορά κατά μήκος των πειραμάτων Π1 και Π2 ανά κατηγορία bul και nobul. Για τη στατιστική ανάλυση επιλέξτε το κατάλληλο μη παραμετρικό τεστ.

Παρουσιάστε τα ευρήματά σας με όσο το δυνατόν αντιπροσωπευτικότερο τρόπο (γραφήματα, confusion matrix (http://en.wikipedia.org/wiki/Confusion_matrix), κλπ) βγάζοντας τα κατάλληλα ειδικά (π.χ., μεταβολή με το κανάλι, είδος ERP, φύλο) και γενικά συμπεράσματα.

Κατεύθυνση 2.- Ανάλυση ERPs #2

a. Δημιουργείστε ένα σήμα μέσου όρου κάνοντας averaging στα σήματα ανά πείραμα (Π1 και Π2) ανά κατάσταση (bul και nobul) και ανά κανάλι, παράγοντας τα αντίστοιχα ERPs για κάθε συμμετέχοντα. Επικεντρωθείτε στις χρονικές περιοχές γύρω από το early posterior positivity (EPN), N170, P300 και late positive potential (LPP) χρησιμοποιώντας κατάλληλο χρονικό παράθυρο. Επιπλέον, παραθυροποιείστε και τα σήματα του smartwatch στα ίδια χρονικά διαστήματα με αυτά του EEG. Στα σήματα που προκύπτουν να υπολογιστεί το πλάτος του Bispectrum (μόνο στην πρωτεύουσα περιοχή). Αφήνεται ελεύθερη η επιλογή των παραμέτρων υπολογισμού του Bispectrum με στόχο την καλύτερη εκτίμησή του. Εάν χρειαστεί κάντε κατάλληλο interpolation (εντολή interp1.m) για να έχετε περισσότερα δείγματα στον υπολογισμό του bispectrum.

b. Για κάθε Bispectrum που υπολογίσατε βρείτε

- (i) την τιμή του μεγίστου του Bispectrum
- (ii) την τιμή των (f1,f2) που αντιστοιχούν στο μέγιστο του Bispectrum
- (iii) το std των συχνοτήτων γύρω από την κορυφή του μεγίστου του Bispectrum (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από το μέγιστο που ξεπερνούν το 10% της τιμής του μεγίστου)
- (iv) την τιμή της επόμενης κορυφής (αμέσως μικρότερη της μέγιστης) του Bispectrum
- (v) την τιμή των (f1,f2) που αντιστοιχούν στην τιμή της επόμενης κορυφής του Bispectrum
- (vi) το std των συχνοτήτων γύρω από την επόμενη κορυφή του μεγίστου του Bispectrum (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από την επόμενη κορυφή που ξεπερνούν το 10% της τιμής της επόμενης κορυφής).



c. Φτιάξτε και το κατάλληλο διάνυσμα χαρακτηριστικών που να αποτελείται από τις τιμές που βρήκατε στο b (δηλ. (i)-(vi)), για κάθε ένα συμμετέχοντα ανά κανάλι και ανά πείραμα και ανά κλάση (P1-bul/nobul, P2-bul/nobul) για να τα χρησιμοποιείτε ως δεδομένα στον ταξινομητή του επόμενου ερωτήματος.

d. Εφαρμόζοντας k-Nearest-Neighbor (k-NN) classification analysis (<http://www.mathworks.com/help/stats/classificationknnnclass.html>) κρατήστε το 75% των δεδομένων για training και το 25% των δεδομένων για testing με τυχαίο τρόπο 5 φορές και βγάλτε τη μέση ακρίβεια στην εκτίμηση των δύο κλάσεων (bul/nobul) ανά πείραμα (P1, P2) για **κάθε ένα συμμετέχοντα ξεχωριστά (user dependent case)**. Σχολιάστε την εκτίμησή σας για το ποια κατηγορία χαρακτηριστικών παρέχει τη μέγιστη δυνατότητα διαφοροποίησης μεταξύ των δύο κλάσεων.

e. Επαναλάβετε τα βήματα c-d, λαμβάνοντας υπόψη την επίδραση του καναλιού, και με βάση τη βιβλιογραφία, επιλέξτε τα κανάλια που μπορεί να σας αυξήσουν την απόδοση της ταξινόμησης. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

f. Επαναλάβετε τα βήματα c-e, λαμβάνοντας υπόψη την επίδραση του φύλου στα δεδομένα σας. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

g. Συγκρίνετε τα αποτελέσματά σας ανά επιλεγμένο ERP και περιγράψτε τυχόν συσχετίσεις, αλληλεπιδράσεις, αλληλουχίες (π.χ., εάν επηρεάζει κάποιο ERP το επόμενο λόγω της χρονικής αλληλουχίας τους).

h. Εξετάστε εάν υπάρχουν στατιστικώς σημαντικές διαφορές στα διανύσματα χαρακτηριστικών του c. κατά μήκος των πειραμάτων P1 και P2 ανά κατηγορία bul και nobul ενοποιώντας τα δεδομένα των χρηστών για το καλύτερο (ή καλύτερα) κανάλι (κανάλια) όπως προκύπτει από το e., για το φύλλο που σας δίνει την καλύτερη ταξινόμηση από το f. και για το ERP που σας δίνει τα καλύτερα αποτελέσματα διαχωρισμού (με βάση το g), ώστε να φανεί η τυχόν επίδραση της εικονικής πραγματικότητας και του βίντεο 360 στη βιωματική αντίληψη του φαινομένου του bullying. Σε δεύτερη προσέγγιση, διαχωρίστε τα δεδομένα από το EEG με αυτά από το smartwatch και δείτε εάν αυτόνομα παρουσιάζουν στατιστικώς σημαντική διαφορά κατά μήκος των πειραμάτων P1 και P2 ανά κατηγορία bul και nobul. Για τη στατιστική ανάλυση επιλέξτε το κατάλληλο μη παραμετρικό τεστ.

Παρουσιάστε τα ευρήματά σας με όσο το δυνατόν αντιπροσωπευτικότερο τρόπο (γραφήματα, confusion matrix (http://en.wikipedia.org/wiki/Confusion_matrix), κλπ) βγάζοντας τα κατάλληλα ειδικά (π.χ., μεταβολή με το κανάλι, είδος ERP, φύλο) και γενικά συμπεράσματα.

Κατεύθυνση 3.- Ανάλυση ERPs #3

a. Δημιουργείτε ένα σήμα μέσου όρου κάνοντας averaging στα σήματα ανά πείραμα (P1 και P2) ανά κατάσταση (bul και nobul) και ανά κανάλι, παράγοντας τα αντίστοιχα ERPs για κάθε συμμετέχοντα. Επικεντρωθείτε στις χρονικές περιοχές γύρω από το early posterior positivity (EPN), N170, P300 και late positive potential (LPP) χρησιμοποιώντας κατάλληλο χρονικό παράθυρο. Επιπλέον, παραθυροποιείτε και τα σήματα του smartwatch στα ίδια χρονικά διαστήματα με αυτά του EEG. Στα σήματα που προκύπτουν να υπολογιστεί ο συνεχής μετασχηματισμός κυματιδίων (continuous wavelet transform-CWT)



χρησιμοποιώντας την εντολή του Matlab `cwt.m` (<http://www.mathworks.com/help/wavelet/ref/cwt.html?searchHighlight=cwt>). Αφήνεται ελεύθερη η επιλογή των παραμέτρων υπολογισμού του `cwt` με στόχο την καλύτερη εκτίμησή του. Δοκιμάστε διαφορετικά mother wavelets και αιτιολογείστε, με βάση τα αποτελέσματα, την τελική επιλογή σας. Μετατρέψτε τα scales σε συχνότητες μέσω της εντολής του Matlab `scal2frq.m` (<http://www.mathworks.com/help/wavelet/ref/scal2frq.html>).

- b. Για κάθε CWT που υπολογίσατε βρείτε
- (i) την τιμή του μεγίστου του CWT
 - (ii) την τιμή της συχνότητας που αντιστοιχεί στο μέγιστο του CWT
 - (iii) το εύρος των συχνοτήτων γύρω από την κορυφή του μεγίστου του CWT (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από το μέγιστο που ξεπερνούν το 10% της τιμής του μεγίστου)
 - (iv) την τιμή της επόμενης κορυφής (αμέσως μικρότερη της μέγιστης) του CWT
 - (v) την τιμή της συχνότητας που αντιστοιχεί στην τιμή της επόμενης κορυφής του CWT
 - (vi) το εύρος των συχνοτήτων γύρω από την επόμενη κορυφή του μεγίστου του CWT (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από την επόμενη κορυφή που ξεπερνούν το 10% της τιμής της επόμενης κορυφής)
- c. Φτιάξτε και το κατάλληλο διάγραμμα χαρακτηριστικών που να αποτελείται από τις τιμές που βρήκατε στο b (δηλ. (i)-(vi)), για κάθε ένα συμμετέχοντα ανά κανάλι και ανά πείραμα και ανά κλάση (P1-bul/nobul, P2-bul/nobul) για να τα χρησιμοποιείτε ως δεδομένα στον ταξινομητή του επόμενου ερωτήματος.
- d. Εφαρμόζοντας Support Vector Machines (SVM) χρησιμοποιώντας την εντολή του Matlab `svmclassify.m` (<http://www.mathworks.com/help/stats/svmclassify.html?searchHighlight=svm>) (υιοθετείστε μια σειρά από SVMs ανά δύο κλάσεις για να λύσετε καλύτερα το πρόβλημα) κρατήστε το 75% των δεδομένων για training και το 25% των δεδομένων για testing με τυχαίο τρόπο 10 φορές και βγάλτε τη μέση ακρίβεια στην εκτίμηση των δύο κλάσεων (bul/nobul) ανά πείραμα (P1, P2) για **ανεξάρτητα του συμμετέχοντα (user independent case)**. Σχολιάστε την εκτίμησή σας για το ποια κατηγορία χαρακτηριστικών παρέχει τη μέγιστη δυνατότητα διαφοροποίησης μεταξύ των δύο κλάσεων.
- e. Επαναλάβετε τα βήματα c-d, λαμβάνοντας υπόψη την επίδραση του καναλιού, και με βάση τη βιβλιογραφία, επιλέξτε τα κανάλια που μπορεί να σας αυξήσουν την απόδοση της ταξινόμησης. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.
- f. Επαναλάβετε τα βήματα c-e, λαμβάνοντας υπόψη την επίδραση του φύλου στα δεδομένα σας. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.
- g. Συγκρίνετε τα αποτελέσματά σας ανά επιλεγμένο ERP και περιγράψτε τυχόν συσχετίσεις, αλληλεπιδράσεις, αλληλουχίες (π.χ., εάν επηρεάζει κάποιο ERP το επόμενο λόγω της χρονικής αλληλουχίας τους).
- h. Εξετάστε εάν υπάρχουν στατιστικώς σημαντικές διαφορές στα διανύσματα χαρακτηριστικών του c. κατά μήκος των πειραμάτων P1 και P2 ανά κατηγορία bul και nobul ενοποιώντας τα δεδομένα των χρηστών για το καλύτερο (ή καλύτερα) κανάλι (κανάλια) όπως προκύπτει από το e., για το φύλλο που σας δίνει την καλύτερη ταξινόμηση από το f. και για το ERP



που σας δίνει τα καλύτερα αποτελέσματα διαχωρισμού (με βάση το g), ώστε να φανεί η τυχόν επίδραση της εικονικής πραγματικότητας και του βίντεο 360 στη βιωματική αντίληψη του φαινομένου του bullying. Σε δεύτερη προσέγγιση, διαχωρίστε τα δεδομένα από το EEG με αυτά από το smartwatch και δείτε εάν αυτόνομα παρουσιάζουν στατιστικώς σημαντική διαφορά κατά μήκος των πειραμάτων Π1 και Π2 ανά κατηγορία bul και nobul. Για τη στατιστική ανάλυση επιλέξτε το κατάλληλο μη παραμετρικό τεστ.

Παρουσιάστε τα ευρήματά σας με όσο το δυνατόν αντιπροσωπευτικότερο τρόπο (γραφήματα, confusion matrix (http://en.wikipedia.org/wiki/Confusion_matrix), κλπ) βγάζοντας τα κατάλληλα ειδικά (π.χ., μεταβολή με το κανάλι, είδος ERP, φύλο) και γενικά συμπεράσματα.

Κατεύθυνση 4.- Ανάλυση ERPs #4

a. Δημιουργείστε ένα σήμα μέσου όρου κάνοντας averaging στα σήματα ανά πείραμα (Π1 και Π2) ανά κατάσταση (bul και nobul) και ανά κανάλι, παράγοντας τα αντίστοιχα ERPs για κάθε συμμετέχοντα. Επικεντρωθείτε στις χρονικές περιοχές γύρω από το early posterior positivity (EPN), N170, P300 και late positive potential (LPP) χρησιμοποιώντας κατάλληλο χρονικό παράθυρο. Επιπλέον, παραθυροποιείστε και τα σήματα του smartwatch στα ίδια χρονικά διαστήματα με αυτά του EEG. Στα σήματα που προέκυψαν να υπολογιστεί ο συνεχής μετασχηματισμός κυματιδίων (continuous wavelet transform-CWT) χρησιμοποιώντας την εντολή του Matlab `cwt.m` (<http://www.mathworks.com/help/wavelet/ref/cwt.html?searchHighlight=cwt>). Αφήνεται ελεύθερη η επιλογή των παραμέτρων υπολογισμού του cwt με στόχο την καλύτερη εκτίμησή του. Δοκιμάστε διαφορετικά mother wavelets και αιτιολογείστε, με βάση τα αποτελέσματα, την τελική επιλογή σας. Μετατρέψτε τα scales σε συχνότητες μέσω της εντολής του Matlab `scal2frq.m` (<http://www.mathworks.com/help/wavelet/ref/scal2frq.html>).

b. Για κάθε CWT που υπολογίσατε βρείτε

- (vi) την τιμή του μεγίστου του CWT
- (vii) την τιμή της συχνότητας που αντιστοιχεί στο μέγιστο του CWT
- (viii) το εύρος των συχνοτήτων γύρω από την κορυφή του μεγίστου του CWT (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από το μέγιστο που ξεπερνούν το 10% της τιμής του μεγίστου)
- (ix) την τιμή της επόμενης κορυφής (αμέσως μικρότερη της μέγιστης) του CWT
- (x) την τιμή της συχνότητας που αντιστοιχεί στην τιμή της επόμενης κορυφής του CWT
- (vii) το εύρος των συχνοτήτων γύρω από την επόμενη κορυφή του μεγίστου του CWT (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από την επόμενη κορυφή που ξεπερνούν το 10% της τιμής της επόμενης κορυφής)

c. Φτιάξτε και το κατάλληλο διάνυσμα χαρακτηριστικών που να αποτελείται από τις τιμές που βρήκατε στο b (δηλ. (i)-(vi)), για κάθε ένα συμμετέχοντα ανά κανάλι και ανά πείραμα και ανά κλάση (Π1-bul/nobul, Π2-bul/nobul) για να τα χρησιμοποιείτε ως δεδομένα στον ταξινομητή του επόμενου ερωτήματος.

d. Εφαρμόζοντας k-Nearest-Neighbor (k-NN) classification analysis (<http://www.mathworks.com/help/stats/classificationknn.html>) κρατήστε το 75% των δεδομένων για training και το 25% των δεδομένων για testing με τυχαίο τρόπο 5 φορές και βγάλτε τη μέση ακρίβεια στην εκτίμηση των δύο κλάσεων (bul/nobul) ανά



πείραμα (Π1, Π2) για **κάθε ένα συμμετέχοντα ξεχωριστά (user dependent case)**. Σχολιάστε την εκτίμησή σας για το ποια κατηγορία χαρακτηριστικών παρέχει τη μέγιστη δυνατότητα διαφοροποίησης μεταξύ των δύο κλάσεων.

e. Επαναλάβετε τα βήματα c-d, λαμβάνοντας υπόψη την επίδραση του καναλιού, και με βάση τη βιβλιογραφία, επιλέξτε τα κανάλια που μπορεί να σας αυξήσουν την απόδοση της ταξινόμησης. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

f. Επαναλάβετε τα βήματα c-e, λαμβάνοντας υπόψη την επίδραση του φύλου στα δεδομένα σας. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

g. Συγκρίνετε τα αποτελέσματά σας ανά επιλεγμένο ERP και περιγράψτε τυχόν συσχετίσεις, αλληλεπιδράσεις, αλληλουχίες (π.χ., εάν επηρεάζει κάποιο ERP το επόμενο λόγω της χρονικής αλληλουχίας τους).

h. Εξετάστε εάν υπάρχουν στατιστικώς σημαντικές διαφορές στα διανύσματα χαρακτηριστικών του c. κατά μήκος των πειραμάτων Π1 και Π2 ανά κατηγορία bul και nobul ενοποιώντας τα δεδομένα των χρηστών για το καλύτερο (ή καλύτερα) κανάλι (κανάλια) όπως προκύπτει από το e., για το φύλλο που σας δίνει την καλύτερη ταξινόμηση από το f. και για το ERP που σας δίνει τα καλύτερα αποτελέσματα διαχωρισμού (με βάση το g), ώστε να φανεί η τυχόν επίδραση της εικονικής πραγματικότητας και του βίντεο 360 στη βιωματική αντίληψη του φαινομένου του bullying. Σε δεύτερη προσέγγιση, διαχωρίστε τα δεδομένα από το EEG με αυτά από το smartwatch και δείτε εάν αυτόνομα παρουσιάζουν στατιστικώς σημαντική διαφορά κατά μήκος των πειραμάτων Π1 και Π2 ανά κατηγορία bul και nobul. Για τη στατιστική ανάλυση επιλέξτε το κατάλληλο μη παραμετρικό τεστ.

Παρουσιάστε τα ευρήματά σας με όσο το δυνατόν αντιπροσωπευτικότερο τρόπο (γραφήματα, confusion matrix (http://en.wikipedia.org/wiki/Confusion_matrix), κλπ) βγάζοντας τα κατάλληλα ειδικά (π.χ., μεταβολή με το κανάλι, είδος ERP, φύλο) και γενικά συμπεράσματα.

Κατεύθυνση 5.- Ανάλυση ERPs #5

a. Δημιουργείτε ένα σήμα μέσου όρου κάνοντας averaging στα σήματα ανά πείραμα (Π1 και Π2) ανά κατάσταση (bul και nobul) και ανά κανάλι, παράγοντας τα αντίστοιχα ERPs για κάθε συμμετέχοντα. Επικεντρωθείτε στις χρονικές περιοχές γύρω από το early posterior positivity (EPN), N170, P300 και late positive potential (LPP) χρησιμοποιώντας κατάλληλο χρονικό παράθυρο. Επιπλέον, παραθυροποιείτε και τα σήματα του smartwatch στα ίδια χρονικά διαστήματα με αυτά του EEG. Στα σήματα που προέκυψαν να εφαρμοστεί Εμπειρική Αποσύνθεση Ρυθμών (Empirical Mode Decomposition) και να υπολογιστούν οι αντίστοιχες Συναρτήσεις Ενδογενών Ρυθμών (Intrinsic Mode Functions-IMFs) και στη συνέχεια ο Hilbert-Huang Transform (HHT) (https://en.wikipedia.org/wiki/Hilbert%E2%80%93Huang_transform, <http://www.mathworks.com/matlabcentral/fileexchange/19681-hilbert-huang-transform>)

b. Για κάθε HHT που υπολογίσατε βρείτε



- (xi) την τιμή του μεγίστου του HHT
- (xii) την τιμή της συχνότητας που αντιστοιχεί στο μέγιστο του HHT
- (xiii) το εύρος των συχνοτήτων γύρω από την κορυφή του μεγίστου του HHT (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από το μέγιστο που ξεπερνούν το 10% της τιμής του μεγίστου)
- (xiv) την τιμή της επόμενης κορυφής (αμέσως μικρότερη της μέγιστης) του HHT
- (xv) την τιμή της συχνότητας που αντιστοιχεί στην τιμή της επόμενης κορυφής του HHT
- (viii) το εύρος των συχνοτήτων γύρω από την επόμενη κορυφή του μεγίστου του HHT (δηλαδή την τυπική απόκλιση της κατανομής των συχνοτήτων που αντιστοιχούν στις τιμές γύρω από την επόμενη κορυφή που ξεπερνούν το 10% της τιμής της επόμενης κορυφής)

c. Φτιάξτε και το κατάλληλο διάγραμμα χαρακτηριστικών που να αποτελείται από τις τιμές που βρήκατε στο b (δηλ. (i)-(vi)), για κάθε ένα συμμετέχοντα ανά κανάλι και ανά πείραμα και ανά κλάση (P1-bul/nobul, P2-bul/nobul) για να τα χρησιμοποιείτε ως δεδομένα στον ταξινομητή του επόμενου ερωτήματος.

d. Εφαρμόζοντας Naive Bayes (NB) Classification χρησιμοποιώντας την κατάλληλη εντολή του Matlab (<http://www.mathworks.com/help/stats/naive-bayes-classification-1.html>) κρατήστε το 75% των δεδομένων για training και το 25% των δεδομένων για testing με τυχαίο τρόπο 10 φορές και βγάλτε τη μέση ακρίβεια στην εκτίμηση των δύο κλάσεων (bul/nobul) ανά πείραμα (P1, P2) **ανεξάρτητα του συμμετέχοντα (user independent case)**. Σχολιάστε την εκτίμησή σας για το ποια κατηγορία χαρακτηριστικών παρέχει τη μέγιστη δυνατότητα διαφοροποίησης μεταξύ των δύο κλάσεων.

e. Επαναλάβετε τα βήματα c-d, λαμβάνοντας υπόψη την επίδραση του καναλιού, και με βάση τη βιβλιογραφία, επιλέξτε τα κανάλια που μπορεί να σας αυξήσουν την απόδοση της ταξινόμησης. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

f. Επαναλάβετε τα βήματα c-e, λαμβάνοντας υπόψη την επίδραση του φύλου στα δεδομένα σας. Συγκρίνεται τα αποτελέσματα της διαδικασίας της ταξινόμησης (βήμα d) με αυτά που προκύπτουν με τη νέα κατηγοριοποίηση και σχολιάστε τα κατάλληλά.

g. Συγκρίνετε τα αποτελέσματά σας ανά επιλεγμένο ERP και περιγράψτε τυχόν συσχετίσεις, αλληλεπιδράσεις, αλληλουχίες (π.χ., εάν επηρεάζει κάποιο ERP το επόμενο λόγω της χρονικής αλληλουχίας τους).

h. Εξετάστε εάν υπάρχουν στατιστικώς σημαντικές διαφορές στα διανύσματα χαρακτηριστικών του c. κατά μήκος των πειραμάτων P1 και P2 ανά κατηγορία bul και nobul ενοποιώντας τα δεδομένα των χρηστών για το καλύτερο (ή καλύτερα) κανάλι (κανάλια) όπως προκύπτει από το e., για το φύλλο που σας δίνει την καλύτερη ταξινόμηση από το f. και για το ERP που σας δίνει τα καλύτερα αποτελέσματα διαχωρισμού (με βάση το g), ώστε να φανεί η τυχόν επίδραση της εικονικής πραγματικότητας και του βίντεο 360 στη βιωματική αντίληψη του φαινομένου του bullying. Σε δεύτερη προσέγγιση, διαχωρίστε τα δεδομένα από το EEG με αυτά από το smartwatch και δείτε εάν αυτόνομα παρουσιάζουν στατιστικώς σημαντική διαφορά κατά μήκος των πειραμάτων P1 και P2 ανά κατηγορία bul και nobul. Για τη στατιστική ανάλυση επιλέξτε το κατάλληλο μη παραμετρικό τεστ.

**Λεόντιος Χατζηλεοντιάδης**

Καθηγητής

Δρ. Χρύσα Παπαδανιήλ, Γιώργος Αποστολίδης

Τ: 2310 996340 F: 2310 996312 E: leontios@auth.grΚτίριο: Δ-6^{ος} όροφος, Γραφείο 3

Παρουσιάστε τα ευρήματά σας με όσο το δυνατόν αντιπροσωπευτικότερο τρόπο (γραφήματα, confusion matrix (http://en.wikipedia.org/wiki/Confusion_matrix), κλπ) βγάζοντας τα κατάλληλα ειδικά (π.χ., μεταβολή με το κανάλι, είδος ERP, φύλο) και γενικά συμπεράσματα.

Εξέταση

Η εξέταση του μαθήματος θα γίνει με παρουσίαση 20 λεπτών από κάθε ομάδα και εξέταση με ερωτήσεις και βαθμολόγηση από τους υπόλοιπους (συμμετοχή στον τελικό βαθμό 50%). Την προηγούμενη ημέρα μπορείτε να στείλετε μέσω e-mail (leontios@auth.gr) όλο το σχετικό υλικό του μαθήματος (αυτή και όσες προηγούμενες εργασίες κάνατε σε μορφή .docx, pptx και τον αντίστοιχο κώδικα).

Μέχρι τις 5.9.2016 θα πρέπει να δηλώσετε τις ομάδες στον ΜΦ Γιώργο Αποστολίδη (12:00 στον 6^ο όροφο)-2310996319 (gkaposto@auth.gr). Θα βγει ανακοίνωση ως προς την ακριβή ημερομηνία εξέτασης (κατά πάσα πιθανότητα 30.9.2016).