



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

Ημερομηνία:

Αριθμ. Πρωτοκόλλου:

ΑΝΑΛΥΤΙΚΟ ΥΠΟΜΝΗΜΑ ΠΡΟΟΔΟΥ ΥΠΟΨΗΦΙΑΣ ΔΙΔΑΚΤΟΡΟΣ

(περίοδος αναφοράς 7-5-2020/30-11-2021)

Όνομα/ Επώνυμο: **Μαρία-Νικολέττα Κολιάρáκη**

Α.Μ. Υ.Δ.: **ΔΜΒ1907**

Ημερομηνία αποδοχής αίτησης από τη συνέλευση του Τμήματος/ Πράξη Συνέλευσης:
7-5-20, Πράξη συνέλευσης: 5/7-5-2020

Ημερομηνία ορισμού 3μελούς συμβουλευτικής επιτροπής & θέματος ΔΔ :
7-5-20, Πράξη συνέλευσης: 5/7-5-2020

Προβλεπόμενο έτος ολοκλήρωσης Δ.Δ.:
2024

Μέλη Τριμελούς Συμβουλευτικής Επιτροπής:

Επιβλέπων: Βεντούρας Ερρίκος-Χαΐμ, Καθηγητής, Τμήμα Μηχανικών Βιοϊατρικής, ΠΑΔΑ

Μέλος: Ασβεστάς Παντελήμων, Αναπληρωτής Καθηγητής, Τμήμα Μηχανικών Βιοϊατρικής, ΠΑΔΑ

Μέλος: Ματσόπουλος Γεώργιος, Καθηγητής, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ, ΕΜΠ

Τίτλος Διδακτορικής Διατριβής (ΔΔ):

«Επεξεργασία Δεδομένων Σακκαδικών Οφθαλμικών Κινήσεων»

1. Περίληψη Αντικειμένου Δ.Δ. (έως 200 λέξεις)

Οι σακκαδικές οφθαλμικές κινήσεις έχουν μελετηθεί τόσο ως δείκτες των νοητικών διεργασιών όσο και ως προς τη σύνδεσή τους με ψυχοπαθολογικούς παράγοντες. Η παρούσα διατριβή αφορά την μελέτη ακολουθιών σακκαδικών κινήσεων των οφθαλμών, κατά τη διάρκεια ενεργούς οπτικής προσήλωσης, με βάση τα χαρακτηριστικά κάθε μιας, με έμφαση στην μελέτη των χρονικών διαστημάτων που

μεσολαβούν ανάμεσα σε δύο διαδοχικές σακκαδικές (Inter-saccade interval – ISI). Σκοπός είναι να διερευνηθεί η ύπαρξη διακριτών καταστάσεων ISIs με την βοήθεια Κρυφών Μαρκοβιανών Μοντέλων (Hidden Markov Model - HMM). Θα χρησιμοποιηθούν δεδομένα οφθαλμικών κινήσεων, που προέρχονται από τη βάση δεδομένων του Αιγινήτειου Νοσοκομείου, η οποία περιέχει σήματα διαφόρων οφθαλμικών διαδικασιών σε 940 άτομα (υγιείς μάρτυρες) σε τρία διαφορετικά πειράματα/δοκιμασίες: (Π1) προσήλωση σε έναν οπτικό στόχο, (Π2) προσήλωση σε ένα οπτικό στόχο ενώ εμφανίζονταν περιφερειακά στόχοι «διάσπασης» σε τυχαία χρονικά διαστήματα και (Π3) προσήλωση χωρίς στόχο (σε λευκή οθόνη). Σε επόμενη φάση θα διερευνηθούν τα μοντέλα HMM που αντιπροσωπεύουν βέλτιστα τις υποκείμενες διαδικασίες δημιουργίας σακκαδικών σε ομάδα ασθενών που έπασχαν από σχιζοφρένεια και ασθενών που έπασχαν από ιδεοψυχαναγκαστική διαταραχή (OCD). Επιδιώκεται η διαλεύκανση των υποκείμενων μηχανισμών διαφοροποίησης των καταστάσεων των ISIs, αλλά και άλλων παραμέτρων των σακκαδικών καθώς και του μηχανισμού δημιουργίας τους σε κάθε πείραμα. Επιπλέον, βασιζόμενοι σε μακροσκοπικές παραμέτρους, θα γίνει προσπάθεια αυτόματης αναγνώρισης της κατηγορίας εξεταζόμενου (υγιείς, σχιζοφρενείς, OCD).

2. Περιγραφή προόδου¹

Κατά τη διάρκεια από την έναρξη της εκπόνησης της διατριβής έως την ημερομηνία του παρόντος υπομνήματος, ασχολήθηκα με την εφαρμογή αλγορίθμων για την δημιουργία Hidden Markov μοντέλων στα δεδομένα υγιών μαρτύρων και την αξιολόγησή τους με βάση τη δυνατότητα αναγνώρισης της δοκιμασίας (Π1, Π2, Π3). Συγκεκριμένα ασχολήθηκα με:

- α. Εξαγωγή χαρακτηριστικών σακκαδικών και ομαδοποίηση ανάλογα με το πλήθος σακκαδικών.
- β. Αναζήτηση του βέλτιστου αριθμού καταστάσεων με χρήση αλγορίθμου Baum-Welch (BW).
- γ. Υπολογισμό Hidden Markov μοντέλων με discriminative training («διακριτική εκπαίδευση» ή «εκπαίδευση διάκρισης»).

A. Εξαγωγή χαρακτηριστικών σακκαδικών και ομαδοποίηση ανάλογα με το πλήθος σακκαδικών

Χρησιμοποιήθηκαν αποθρομβωποιημένα δεδομένα, τα οποία έχουν χρησιμοποιηθεί σε προηγούμενη μελέτη [Korda et al, 2016] με χρήση Semi-Markov μοντέλων. Από αυτά δημιουργήθηκαν αρχεία δεδομένων που περιλαμβάνουν για κάθε σακκαδική: διάρκεια σακκαδικής (duration), ISI, πλάτος (angular deviation), μέγιστη ταχύτητα (maximum velocity), μέση ταχύτητα (mean velocity), απόλυτη τιμή πλάτους, απόλυτη τιμή μέγιστης ταχύτητας και απόλυτη τιμή μέσης ταχύτητας. Συνολικά, 742 σήματα από την δοκιμασία Π1 (προσήλωση σε οπτικό στόχο), 789 σήματα από την δοκιμασία Π2 (προσήλωση σε οπτικό στόχο ενώ εμφανίζονται περιφερειακά στόχοι «διάσπασης» σε τυχαία χρονικά διαστήματα) και 554 σήματα

¹ Σύμφωνα με τον Κανονισμό Διδακτορικών Σπουδών του Τμήματος «Ο ΥΔ κάθε έτος, στο διάστημα από το Σεπτέμβριο έως το Νοέμβριο, παρουσιάζει προφορικά και υποβάλλει εγγράφως αναλυτικό υπόμνημα ενώπιον της ΤΣΕ σχετικά με την πρόοδο της διδακτορικής διατριβής. Αντίγραφο του υπομνήματος, καθώς και σχόλια επ' αυτού από τον ΕΚ ή/και την ΤΣΕ καταχωρούνται στον ατομικό φάκελο του ΥΔ». Δεδομένης της έναρξης της Διατριβής τον Μάιο του 2020, το παρόν υπόμνημα είναι το πρώτο, καλύπτοντας έναν πλήρη ημερολογιακό χρόνο και το πρόσθετο διάστημα έως το Νοέμβριο του 2021.

από την δοκιμασία P3 (προσήλωση χωρίς στόχο) ταξινομήθηκαν σύμφωνα με τον αριθμό σακκαδικών. Στη βιβλιογραφία χρησιμοποιούνται σήματα με ίδιο αριθμό μεταβάσεων για κάθε συνολική χρονοσειρά σήματος (ίδιου «μεγέθους»/«πλήθους» μεταβάσεων), για τον λόγο αυτό επιλέγουμε να εφαρμόσουμε τον αλγόριθμο σε σήματα με ίδιο αριθμό σακκαδικών, θεωρώντας ότι κάθε σακκαδική αποτελεί εμφάνιση μετάβασης. Γενικώς, στις δοκιμασίες P1 και P2 πάνω από 85% των σημάτων περιέχει έως 20 σακκαδικές, ενώ στη δοκιμασία P3 το αντίστοιχο ποσοστό είναι 45,3%. Επιλέχθηκε να μελετηθεί, αρχικά, η κατηγορία με 16 σακκαδικές που εμφανίζονται σε 20 σήματα της P1, 23 σήματα της P2 και 21 σήματα της P3.

Αναφορές σε προηγούμενες μελέτες:

Korda A. I., M. Koliaraki, P. A. Asvestas, G. K. Matsopoulos, E. M. Ventouras, P. Y. Ktonas, N. Smyrnis, "Discrete states of attention during active visual fixation revealed by markovian analysis of the time series of intrusive saccades", *Neuroscience*, vol. 339, pp. 385-395, 2016.

B. Αναζήτηση του βέλτιστου αριθμού καταστάσεων με χρήση αλγορίθμου Baum-Welch

Με χρήση εργαλείων HMM δημιουργήσαμε μοντέλα με τη μέθοδο leave-one-out για τις δοκιμασίες P1, P2, P3 για 2, 3 και 4 καταστάσεις P1K2, P1K3, P1K4, P2K2, P2K3, P2K4, P3K2, P3K3, P3K4. Σε κάθε μοντέλο υπολογίστηκαν οι παράμετροι μ , σ , A , χρησιμοποιώντας ως ομάδα εκπαίδευσης (EK) όλα τα σήματα εκτός από ένα (ΕΛ). Κάθε σήμα επιλέχθηκε κυλιόμενα ως ΕΛ σε κάθε μοντέλο και χρησιμοποιήθηκε για τον έλεγχο. Για την εκπαίδευση του μοντέλου χρησιμοποιήθηκε ο αλγόριθμος BW.

Υπολογίζοντας την πιθανοφάνεια (log-likelihood, LL) κάθε μοντέλου κατασκευάστηκε ο πίνακας απόδοσης (δεν πρέπει να συγχέεται με απόδοση ταξινόμησης), όπου βέλτιστο θεωρείται το μοντέλο που επιτυγχάνει ελάχιστη τιμή LL. Προκειμένου να ελεγχθεί η απόδοση του αλγορίθμου δημιουργήθηκαν μοντέλα 1 έως 5 καταστάσεων και εφαρμόστηκαν τα κριτήρια AIC (Akaike Information Criterion) και BIC (Bayesian Information Criterion). Με χρήση των κριτηρίων AIC και BIC και βελτιώσεών τους που αναφέρονται στην βιβλιογραφία και συγκριτική αξιολόγηση των τιμών που δίνουν τα κριτήρια φάνηκε ότι ο «βέλτιστος» αριθμός καταστάσεων, ως προς τις επιδόσεις σύγκλισης του μοντέλου (τιμή LL), προκύπτει για μοντέλο με 3 καταστάσεις σε κάθε δοκιμασία.

Στη συνέχεια, προκειμένου να ελέγξουμε την επίδοση ταξινόμησης των μοντέλων που δημιουργήθηκαν, δοκιμάσαμε για κάθε εξεταζόμενο όλους τους πιθανούς συνδυασμούς αριθμού διακριτών καταστάσεων (δηλ. αντίστοιχων μοντέλων) σε κάθε δοκιμασία 111,112,...,115,...,555. Αυτό έγινε χάριν πληρότητας, ώστε να δοκιμασθούν και μοντέλα που η διαδικασία επιλογής βέλτιστου μοντέλου βάσει των τιμών σύγκλισης του BW είχε αποκλείσει, δεδομένης και της αβεβαιότητας που υπάρχει στην χρήση των κριτηρίων AIC & BIC και των παραλλαγών τους. Τα ποσοστά ταξινόμησης που προέκυψαν δεν κρίθηκαν αποδεκτά. Θεωρήθηκε επομένως αναγκαίο να εφαρμόσουμε μεθόδους discriminative training για HMM.

Γ. Υπολογισμός Hidden Markov μοντέλων με discriminative training

Για τη διερεύνηση μεθόδων discriminative training μελετήθηκαν κυρίως τα άρθρα:

1. Simola J., Salojärvi J. & Kojo I., Using hidden Markov model to uncover processing states from eye movements in information search tasks, Cognitive Systems Research, Vol. 9, Is. 4, pp. 237-251, October 2008
2. Moerland T. (2015), Knowing what you don't know, Master Thesis, Universiteit Leiden, Netherlands

Η μέθοδος discriminative training χρησιμοποιεί τον αλγόριθμο Extended Baum-Welch (EBW). Σε αντίθεση με τον BW, όπου η εκπαίδευση κάθε μοντέλου γίνεται στηριζόμενοι στην βελτιστοποίηση της απόδοσης του αλγόριθμου ταξινόμησης θεωρώντας τα δεδομένα κάθε κατηγορίας (δοκιμασίες) ξεχωριστά (δηλ. μία κατηγορία είναι η κατηγορία «στόχος» και όλες οι άλλες αποτελούν αδιαφοροποίητα την κατηγορία «μη-στόχος»), στον EBW η βελτιστοποίηση της απόδοσης του αλγόριθμου ταξινόμησης γίνεται με τρόπο που βελτιστοποιεί την διάκριση (εξ ου και η ονομασία discriminative training) μεταξύ μοντέλου «στόχου» και καθενός από τα υπόλοιπα «ανταγωνιστικά» πλέον μοντέλα. Για τον υπολογισμό των μοντέλων με EBW χρησιμοποιήθηκε ο δημόσια διαθέσιμος κώδικας του Moerland (2015). Για αυτόν τον σκοπό τα δεδομένα χρειάστηκε να προσαρμοστούν στην μορφή του κώδικα. Δημιουργήθηκε ένας σύνθετος πίνακας με τρεις στήλες: στην πρώτη είναι ένας αριθμός που αντιστοιχεί στη δοκιμασία, στη δεύτερη είναι ο κωδικός του εξεταζόμενου και στην τρίτη το διάλυμα των ISI.

Έγιναν οι κατάλληλες προσαρμογές προκειμένου να επιλεγούν οι εντολές του κώδικα που ταιριάζουν στις ανάγκες των δεδομένων που χρησιμοποιούμε. Δοκιμάστηκαν διαφορετικές τιμές για την παράμετρο U. Η παράμετρος αυτή διαφέρει στον κώδικα σε σχέση με τη θεωρητική παράμετρο U του EBW. Στην παρούσα φάση διεξάγονται προκαταρκτικές δοκιμές της χρήσης του κώδικα στα δεδομένα μας για να διαπιστωθούν τυχόν λάθη στην προσαρμογή που έγινε στον κώδικα, καθώς και η συμπεριφορά του κώδικα και του ίδιου του αλγόριθμου στα δεδομένα μας, τα οποία έχουν δύο εγγενείς αδυναμίες, δηλ. την ύπαρξη μικρού σχετικά αριθμού μεταβάσεων και δειγμάτων και την μη σαφή διαφοροποίηση καταστάσεων, όσον αφορά τουλάχιστον τα ISIs.

3. Σύνοψη νέων αποτελεσμάτων (έως 200 λέξεις)

1. Ολοκληρώθηκε η μελέτη διακριτών καταστάσεων με χρήση Hidden-Markov Model με κώδικα Baum-Welch και προέκυψε ο βέλτιστος αριθμός 3 καταστάσεων σε κάθε δοκιμασία.
2. Έγινε προσαρμογή των δεδομένων και τροποποιήσεις του δημόσια διαθέσιμου αλγόριθμου discriminative training με χρήση Extended Baum Welch. Στην παρούσα φάση δεν έχουμε αποφανθεί ακόμα για το βέλτιστο μοντέλο βάσει αυτής της μεθόδου.

4. Δημοσιεύσεις, συμμετοχή σε συνέδρια και ημερίδες

Δημοσιεύσεις σε Επιστημονικά Περιοδικά: -

Δημοσιεύσεις σε Επιστημονικά Συνέδρια: -

Υποβληθείσες εργασίες σε Επιστημονικά Περιοδικά: -

Υποβληθείσες Εργασίες σε Επιστημονικά Συνέδρια: -

Άλλες Εργασίες/Ανακοινώσεις/Παρουσιάσεις/ Λοιπό Έργο: -

5. Σύντομη αναφορά στον χρονικό προγραμματισμό και πιθανές καθυστερήσεις (έως 200 λέξεις)

Στο επόμενο έτος 2022, στόχοι της έρευνας είναι:

- (α) Η χρήση συνθετικών δεδομένων, με σαφώς διακριτές καταστάσεις μεταβάσεων, ώστε να μελετηθεί με τρόπο ελεγχόμενο ως προς τα χαρακτηριστικά των δεδομένων, η λειτουργία τόσο του EBW, όσο και του BW, λαμβάνοντας υπόψη τις έως τώρα ασάθειες που παρατηρούμε στην λειτουργία του EBW στα δεδομένα μας, αλλά και την αποτυχία του BW να δώσει αποδεκτά ποσοστά ταξινόμησης στα δεδομένα μας. Η ολοκλήρωση της δοκιμής των δύο μεθόδων με τα συνθετικά δεδομένα αναμένεται να δώσει ξεκάθαρη εικόνα για τα πλεονεκτήματα της μεθόδου discriminative training σε σχέση με το απλό Baum-Welch.
- (β) Η συνέχιση και ολοκλήρωση της μελέτης των Hidden Markov μοντέλων με discriminative training, ώστε: i) να γίνει διαχωρισμός των διακριτών καταστάσεων και ii) να διαλευκανθεί η διαφοροποίηση του μηχανισμού προσήλωσης στην κάθε δοκιμασία.

Υπογραφή

Μαρία-Νικολέττα Κολιαράκη