



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΙΑΤΡΙΚΗΣ

“Alternative medical data structures to support fast and effective medical decision making”

Επιβλέπουσα καθηγήτρια: Δρ. Καλλέργη Μαρία
Συμβουλευτική επιτροπή: Δρ. Γκλώτσος Δημήτριος, Δρ. Μιχαήλ Χρήστος

Υποψήφιος Διδάκτωρ: Μαρίνης Μιχαήλ

Δεκέμβριος 2023

Ο σκοπός της διατριβής είναι να μπορέσουμε να δημιουργήσουμε ένα πρότυπο με το οποίο θεωρητικά θα μπορούν να αξιοποιηθούν οι σύγχρονες τεχνολογίες με τρόπο που να εξάγουμε ασφαλή γρήγορα και συνδυαστικά συμπεράσματα από τον ιατρικό φάκελο των ασθενών για να αξιοποιηθούν στην αντιμετώπιση της ασθένειας άλλων ασθενών.

Η ιδέα αναλύεται ως εξής. Ο κάθε ασθενής αποτελεί ένα δυναμικό σύστημα μεταβλητό στο χρόνο το οποίο μεταβάλλεται προς μια συγκεκριμένη κατεύθυνση θεραπείας ή επιδείνωσης. Η σημασία της μεταβολής στο χρόνο θεωρούμε πως είναι τεράστια με δεδομένο ότι η συμπτωματολογία μεταβάλλεται αναλόγως της χρονικής στιγμής. Αν λοιπόν για παράδειγμα κάποιος ασθενής εμφανίζει υψηλό πυρετό και του χορηγηθεί αντιπυρετικό τότε ο πυρετός παύει να δίνει ένδειξη που μπορεί να χρησιμοποιηθεί ως χαρακτηριστικό ταξινόμησης ή ανάλυσης των δεδομένων του ασθενούς. Αν σε αυτόν τον ασθενή διακόψουμε τη χρονική στιγμή T1 τη χορήγηση αντιπυρετικού οι πιθανότητες να εμφανίζει ξανά πυρετό μπορεί να είναι εντελώς διαφορετικές σε σχέση με το να διακόψουμε τη χορήγηση σε χρόνο T2.

Οι κλινικές δοκιμές που έχουμε στη διάθεσή μας είναι εντελώς στοχευμένες και με επιστημονικό τρόπο μας δίνουν οδηγίες για το πώς πρέπει να αντιμετωπιστεί μια ασθένεια με χορηγήσεις ή όχι φαρμάκων ή θεραπευτικών αγωγών και σε ποιο βάθος χρόνου ώστε να επιτευχθεί θεραπεία. Ωστόσο αυτό που θεωρούμε ως ιδέα είναι ότι εάν δομηθεί κατάλληλα ο ιατρικός φάκελος και αξιοποιεί σε πραγματικό χρόνο τα δεδομένα εξέλιξης της κατάστασης ενός ασθενούς τότε η ίδια η αναφορά στο ιστορικό του ασθενούς θα μπορεί να αποτελεί σημείο έρευνας μιας κλινικής δοκιμής. Αν συνδυαστούν πολλοί φάκελοι από πολλούς ασθενείς τότε δημιουργείται μια δεξαμενή δεδομένων από την οποία είναι πιθανό να μπορούμε να εξάγουμε χρήσιμα συμπεράσματα για να βοηθήσουμε την θεραπεία άλλων ασθενών με παρόμοια συμπτωματολογία, ηλικία, φύλο κλπ. Όσο μεγαλύτερος θα είναι ο όγκος δεδομένων τόσο περισσότερο ασφαλή θα είναι τα συμπεράσματά μας και τόσο καλύτερα θα μπορεί να αντιμετωπιστεί μια νόσος, με δεδομένο ότι η ιατρική είναι απόλυτα συνυφασμένη πλέον με τη στατιστική.

Πορεία - Εργαλεία:

Στο τέλος της διδακτορικής διατριβής επιδιώκουμε να παραδοθεί κώδικας και μια εν λειτουργία πιλοτική πλατφόρμα η οποία θα προσπαθήσουμε να υλοποιεί τους σκοπούς της διατριβής. Για το λόγο αυτό θα χρησιμοποιηθούν τεχνολογίες υπολογιστών οι οποίες θα πρέπει να μπορούν να διαχειριστούν με ασφάλεια μεγάλο όγκο δεδομένων και συνδέσεων. Η εφαρμογή θα πρέπει να χρησιμοποιεί V8 τεχνολογία http server μέσω της οποίας θα μπορούν να γίνουν ασύγχρονες συνδέσεις. Όταν ο όγκος χρηστών πολλαπλασιάζεται πρέπει οι διαδικασίες να μην απασχολούν τον web server διότι η υπολογιστική δύναμη και η μνήμη του είναι πεπερασμένα μεγέθη. Για αυτό το λόγο οφείλουμε να χρησιμοποιήσουμε τεχνολογίες όπου γίνεται όσο το δυνατό μικρότερη χρήση του web server σε ανάλυση δεδομένων. Για να το πετύχουμε αυτό επιλέξαμε η εφαρμογή να υλοποιηθεί με χρήση node express.js. Για τη βάση δεδομένων αρχικά επιλέχθηκε η πλατφόρμα Cassandra καθώς είναι πολλά υποσχόμενη σύγχρονη πλατφόρμα με εξειδίκευση στα μεγάλα δεδομένα (big data). Ωστόσο η αποσφαλτάωση (debugging) στο περιβάλλον που παρέχει είναι έλλειπης και δύσκολη, οπότε ίσως τελικά καταλήξουμε σε κάποια mongodb ή ακόμα και mysql. Ο server θα λειτουργεί σε debian linux και ο προγραμματισμός του θα γίνεται με sshfs και ssh συνδέσεις.

Στη συνέχεια αρχίσαμε να ψάχνουμε ποιες θεωρίες μπορεί να αναλύουν τα δυναμικά συστήματα με τρόπο που να είναι δόκιμος προς την ερευνά μας. Αρχικά αρχίσαμε να μελετάμε το

CMDT (Current Medical Diagnosis Treatment). Σκοπός αυτού ήταν η καλύτερη και σε βάθος κατανόηση του πώς λαμβάνεται η ιατρική απόφαση. Σε αυτό το βιβλίο βρήκαμε ότι οι αποφάσεις ενός κλινικού ιατρού είναι τεκμηριωμένες πάνω σε κλινικές δοκιμές και σε EBM (evidence based medicine). Στη συνέχεια λοιπόν αρχίσαμε τη μελέτη του Randomized Clinical Trial Design. Σε αυτό το βιβλίο μας δόθηκαν πληροφορίες για το πώς στήνεται μια κλινική δοκιμή και με ποιον τρόπο διασφαλίζεται η αξιοπιστία της. Ωστόσο δεν μας δόθηκε σημαντική πληροφορία που να σχετίζει με χρονικά μεταβαλλόμενο δυναμικό τρόπο κάποια προτυποποίηση κλινικών δοκιμών. Επίσης αρκετές ορολογίες ήταν ορολογίες καθαρής ιατρικής στατιστικής οπότε έπρεπε να δούμε ξανά μέσα στην βιοστατιστική εάν υπάρχει κάτι χρήσιμο για το σκοπό μας. Σε αυτή την αναζήτηση βρήκαμε αναγκαία την χρήση νέων εργαλείων καθώς και έτοιμων μοντέλων και μεθόδων με τη χρήση γλώσσας R. Έγινε μια προσπάθεια να εξαχθούν δεδομένα που να μπορούν να βοηθήσουν στην ερευνά μας αλλά πάλι υπήρχε πρόβλημα στον τρόπο τυποποίησης του χρόνου ως παράμετρο στην κλινική δοκιμή. Ταυτόχρονα αρχίσαμε να ψάχνουμε λύση θεωρητική σε διάφορα άλλα επίπεδα θεωριών που έχουν το δυναμικό χαρακτήρα του χρόνου ενσωματωμένο στη θεωρία τους. Ψάξαμε για πιθανή εφαρμογή κυτταρικών αυτομάτων. Δεν βρήκαμε ακόμα ότι μπορεί κάπως να είναι χρήσιμη. Θεωρήσαμε ότι η εξέλιξη μιας νόσου είναι χασοτική και επομένως ίσως έχει νόημα να ανιχνευθεί μια εργοδικότητα άλλα επίσης δεν μπορέσαμε να τυποποιήσουμε τον τρόπο. Είδαμε τη χρήση μοντέλων Bayesian και επιλέξαμε να την επανεκτιμήσουμε θετικά προς τον σκοπό μας. Μελετήσαμε την περίπτωση της πολυμεταβλητής στατιστικής ανάλυσης και είδαμε ότι υπάρχει μια δυσκολία εφαρμογής στον τρόπο που ορίζουμε τις σταθερές του δυναμικού μας συστήματος αλλά επίσης το μελετούμε. Προσπαθήσαμε να δούμε και άλλες μεθόδους εξομοίωσης πολύπλοκων μεταβλητών συστημάτων και σε αυτή τη φάση προσπαθούμε να δούμε αν η χρήση transformers με AI μπορεί να μας δομηθεί ώστε να μας δώσει τρόπους που να εξαγάγουμε χρήσιμα συμπεράσματα.

1. Ταυτόχρονα αρχίσαμε να κοιτάμε την εξέλιξη και την έρευνα στον τομέα των ηλεκτρονικών ιατρικών δεδομένων. Μελετήσαμε το βιβλίο *New Perspectives on medical Records* όπου και βρέθηκε ότι υπάρχουν διάφορα προβλήματα “διαχρονικά” Η δυσκολία που συναντάμε εκφράζεται σε βασικά άλματα προβλήματα. Από αυτά, εκείνα που μας απασχολούν περισσότερο είναι τα παρακάτω: Με ποιο τρόπο μπορούμε να πούμε ότι ένας ασθενής καλυτέρευσε ή επιδεινώθηκε;
2. Πρέπει να εκτιμάται η εικόνα του ασθενούς μετά την έξοδο του από το θεραπευτικό ίδρυμα και με ποιο τρόπο και για πόσο χρόνο;
3. Πώς μπορούμε να ορίσουμε την ελάχιστη “κλινική χρονική μονάδα”; Μονάδα δηλαδή χρόνου που θα μετρά τις μεταβολές της κατάστασης του ασθενούς, (Time Intervals, Constants and Variables definition and modeling). Μπορεί αυτή να είναι γενική ή χρειάζονται διαφορετικές προσεγγίσεις;
4. Πώς μπορούμε να βρούμε δυναμικά δεδομένα μέσω των οποίων να μπορούμε να δημιουργήσουμε πρακτική εφαρμογή; (Databases usefull information)
5. Με ποιο τρόπο μπορεί ένας κλινικός ιατρός να εισάγει δεδομένα που να αφορούν την εξέλιξη ενός στατιστικού συστήματος, ενώ ταυτόχρονα πρέπει ενεργήσει και να πάρει αποφάσεις σοβαρές στο χρόνο εργασίας τους (Data Entry). Ομοίως μπορεί ένας νοσηλευτής να υποστηρίξει στην εισαγωγή δεδομένων ή τελικά χρειάζεται εξεικευμένο προσωπικό και σε τί βαθμό;
6. Με ποιο τρόπο μπορεί ένας ιατρός να δει γρήγορα τα δεδομένα που τον ενδιαφέρουν χωρίς να μπερδευτεί και χωρίς να αποπροσανατολιστεί; (Visualization of data)

Προέκυψε λοιπόν η ανάγκη να ψάξουμε να βρούμε τι υπάρχει από τους επίσημους φορείς , ποιες οδηγίες δίδονται και πώς προχωρά η επίσημη έρευνα και βρήκαμε ότι για τους σκοπούς της διατριβής πρέπει να αξιοποιηθεί η γνωσιακή και ερευνητική βάση κάποιου επίσημου οργανισμού. Στην Ευρωπαϊκή ένωση καταλήξαμε στον οργανισμό TEHDAS <https://tehdas.eu/> Joint Action

Towards the European Health Data Space – TEHDAS ο οποίος ασχολείται ακριβώς με τη μοντελοποίηση των ιατρικών δεδομένων για μελλοντική χρήση σε πανευρωπαϊκό επίπεδο ενός πανευρωπαϊκού ιατρικού φακέλου για τους κατοίκους της Ε.Ε.

Στην παρούσα φάση προσπαθούμε επίσης να κατανοήσουμε τις οδηγίες του εν λόγω οργανισμού και να δούμε πώς μπορεί να αξιοποιηθεί το έργο του στην πρόοδο της διατριβής.

Ταυτόχρονα έγινε μια προσπάθεια ανίχνευσης του πώς μπορούν τα φάρμακα να επιδρούν στην βελτίωση ή επιδείνωση της υγείας ενός ασθενούς με ταυτόχρονες διαφορετικές νόσους αναλόγως με τη σειρά με την οποία χορηγούνται. Σε αυτό το κομμάτι βρήκαμε ότι μοναδικό κριτήριο επί του παρόντος για το πώς επιδρούν μεταξύ τους τα φάρμακα είναι το drug-drug interaction χωρίς να γνωρίζουμε πάντα τη χρονική έκφραση της όποιας αντίδρασης σε συνδυασμούς. Ωστόσο η drug-drug interaction είναι μια πολύ ισχυρά τεκμηριωμένη πρακτική θεωρία που χωρίζει επί της ουσίας τον οργανισμό σε κάποια βασικά λειτουργικά βιολογικά πεδία και ταξινομεί τα φάρμακα αναλόγως με το πώς επιδρούν στα πεδία αυτά. Κατά συνέπεια είναι δύσκολο να ισχυριστούμε ότι κάποιος συνδυασμός φαρμάκων είναι εσφαλμένος από τη στιγμή που δίδεται έμφαση στην αντίδραση μεταξύ των φαρμάκων. Ωστόσο το παραπάνω δεν αναιρεί την πιθανή σημασία που έχει στην αντιμετώπιση ενός περιστατικού με συνοσηρότητες, η επιλογή της χρονικής σειράς αντιμετώπισης των νόσων.

Βρήκαμε επίσης ότι η απεικόνιση των χρονικά μεταβλητών δεδομένων αποτελεί πρόκληση για όποια επιστήμη ασχολείται με τους ιατρικούς φακέλους με δεδομένο ότι δεν είναι ούτε τυποποιημένη ούτε εύκολη. Στην παρούσα φάση απο τον χώρο των γραφικών υπολογιστών γίνονται διάφορες προτάσεις για τον καλύτερο δυνατό τόπο απεικόνισης των χρονικά μεταβλητών περιστατικών (Visual Analytics of Electronic Health Records with a Focus on Time).

Προσπαθώντας να βρούμε πώς γίνεται εκτίμηση για το πότε ένας ασθενής θεωρείται ότι είναι καλύτερα ή χειρότερα ή σε σταθερή κατάσταση βρήκαμε ότι ίσως μια εύκολη γρήγορη απεικόνιση θα ήταν με τη χρήση κάποιας παρόμοιας μεθοδολογίας με την μέθοδο της pictal health με ηλεκτρονικό τρόπο. Ωστόσο δεν είναι ακόμα σαφής ο τρόπος με τον οποία θα μπορούσε να γίνει χρήση της συγκεκριμένης μεθοδολογίας.

Τέλος έγινε προσπάθεια να δούμε εάν έχουμε X ασθένειες και Y σκευάσματα προς χορήγηση στον ίδιο ασθενή, πόσες ερευνητικές πιθανότητες της σειράς χορήγησης μπορεί να υπάρχουν. Διατυπώσαμε έναν μαθηματικό τύπο που χρήζει περεταίρω απόδειξης και σε πρώτη φάση μας δείχνει ότι όσο αυξάνονται τα σκευάσματα και οι παθήσεις η ιδιαιτερότητα της χρονικής σειράς χορήγησης σκευασμάτων (ή θεραπειών) αυξάνεται γεωμετρικά με τρόπο που γίνεται πρακτικά αδύνατη η ιχνηλάτηση της βέλτιστης μεθόδου με αναλυτικό τρόπο.

Στόχοι Στο μέλλον θα προσπαθήσουμε να οργανώσουμε να ταξινομήσουμε και να αξιοποιήσουμε όλα τα παραπάνω με τρόπο που να οδηγούν σε χρήσιμα συμπεράσματα για την ανάπτυξη εργαλείων εφαρμογών.

Ευχαριστώ για το χρόνο σας
Μαρίνης Μιχαήλ