

1. Κατασκευή Βιολογικών Βάσεων Δεδομένων

Με την εκρηκτική αύξηση του ογκου των διαθέσιμων ακολουθιών DNA και πρωτεϊνών, σημαντικό ρόλο στη σύγχρονη Βιοπληροφορική παίζει η ανάπτυξη ολοκληρωμένων βάσεων δεδομένων που περιέχουν πληροφορίες για τις ακολουθίες αυτές. Πέραν των ήδη γνωστών και δημόσια διαθέσιμων βάσεων δεδομένων γενικού περιεχομένου, ολοένα αυξάνεται τα τελευταία χρόνια η ανάγκη για ύπαρξη βάσεων δεδομένων με εξειδικευμένο περιεχόμενο, βάσεις δηλαδή που περιέχουν ακολουθίες μιας συγκεκριμένης κατηγορίας αλλά με υψηλό βαθμό και καλής ποιότητας σχολιασμό που αφορά τη λειτουργία τους. Εκτός από την αναγκαιότητά τους στη μελέτη των βιολογικών συστημάτων, είναι όλο και περισσότερο απαραίτητες για την δημιουργία πρότυπων συνόλων δεδομένων πάνω στα οποία θα εκπαιδευθούν και θα ελεγχθούν (training and test) με αποτελεσματικότητα νέοι αλγόριθμοι Βιοπληροφορικής. Στα πλαίσια αυτής της εργασίας θα γίνει προσπάθεια να κατασκευασθεί μια πρότυπη βάση δεδομένων η οποία θα περιέχει αξιόπιστα σύνολα πρωτεϊνικών ακολουθιών με ειδικό ρόλο, όπως οι πρωτεΐνες εξωτερικής μεμβράνης των μιτοχονδρίων και των χλωροπλαστών. Τα δεδομένα θα συλλεχθούν από τη βιβλιογραφία και θα παρουσιαστούν σε ενιαία μορφή, ενώ θα κατασκευαστεί και διαδικτυακή εφαρμογή.

Επιθυμητά προσόντα: Βασικές γνώσεις βιολογίας και βιοπληροφορικής, γνώσεις σε βάσεις δεδομένων, βασικές γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος

2. Βιολογία συστημάτων και ασθένειες πολυπαραγοντικής αιτιολογίας

Οι ασθένειες πολυπαραγοντικής αιτιολογίας καθορίζονται από την συνύπαρξη και αλληλεπίδραση πολλών διαφορετικών αιτιολογικών παραγόντων, τόσο γενετικών όσο και περιβαλλοντικών. Στην εργασία αυτή θα μελετηθούν με όρους βιολογίας συστημάτων (ανάλυση βιολογικών δικτύων, χρήση βάσεων δεδομένων κ.ο.κ.) οι ασθένειες πολυπαραγοντικής αιτιολογίας (καρκίνος, καρδιαγγειακά κ.ο.κ.) καθώς θα συλλεχθούν και θα μελετηθούν συνοπτικά δεδομένα για αυτές όπως γονίδια που τις επηρεάζουν, κληρονομησιμότητα, παράγοντες κινδύνου, συννοσηρότητες κ.α.

Επιθυμητά προσόντα: Γνώσεις βιολογίας και βιοπληροφορικής, γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος

3. Λογισμικό μέτα-ανάλυσης στο Stata

Το Stata είναι ένα στατιστικό πακέτο με πολύ μεγάλες δυνατότητες ανάλυσης δεδομένων ειδικά στο χώρο της βιοστατιστικής και της επιδημιολογίας. Το πακέτο διαθέτει ενσωματωμένη και τη δική του γλώσσα προγραμματισμού για τη διευκόλυνση σύνθετων αναλύσεων και παρέχει έτσι τη δυνατότητα στο χρήστη να γράψει δικές του μακροεντολές. Στη συγκεκριμένη εργασία θα γίνει εκμάθηση της γλώσσας προγραμματισμού του Stata και θα πραγματοποιηθεί υλοποίηση πρωτοποριακών μεθόδων μέτα-ανάλυσης σε μακροεντολές έτσι ώστε να γίνουν αυτές διαθέσιμες στην επιστημονική κοινότητα. Οι εντολές θα εφαρμοστούν σε πρωτότυπα βιοϊατρικά δεδομένα.

Επιθυμητά προσόντα: Καλές γνώσεις στατιστικής, βασικές γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος

4. Ανάπτυξη Λογισμικού Βιοστατιστικής στο Stata

Το Stata είναι ένα στατιστικό πακέτο με πολύ μεγάλες δυνατότητες ανάλυσης δεδομένων ειδικά στο χώρο της βιοστατιστικής και της επιδημιολογίας. Το πακέτο διαθέτει ενσωματωμένη και τη δική του γλώσσα προγραμματισμού για τη διευκόλυνση σύνθετων αναλύσεων και παρέχει έτσι τη δυνατότητα στο χρήστη να γράψει δικές του μακροεντολές. Στη συγκεκριμένη εργασία θα γίνει εκμάθηση της γλώσσας προγραμματισμού του Stata και θα πραγματοποιηθεί υλοποίηση πρωτοποριακών στατιστικών μεθόδων σε μακροεντολές έτσι ώστε να γίνουν αυτές διαθέσιμες στην επιστημονική κοινότητα. Οι εντολές θα εφαρμοστούν σε πρωτότυπα βιοϊατρικά δεδομένα.

Επιθυμητά προσόντα: Καλές γνώσεις στατιστικής, βασικές γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος

5. Μεθοδολογίες ανάλυσης κειμένου στην Εκπαιδευτική Έρευνα

Οι ποιοτικές μέθοδοι (ανάλυση περιεχομένου, συνεντεύξεις κ.ο.κ.) έχουν σημαντική θέση στην εκπαιδευτική έρευνα. Στην εργασία αυτή θα ασχοληθούμε με υπολογιστικές μεθοδολογίες ανάλυσης κειμένου (text mining), θα συλλεχθούν και θα μελετηθούν δεδομένα από υπάρχουσες μελέτες και θα αναπτυχθούν πρωτόκολλα και λογισμικά χρήσιμα στην επιστημονική κοινότητα.

Επιθυμητά προσόντα: γνώσεις εκπαιδευτικής έρευνας, γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος

6. Κατασκευή διαδικτυακής πλατφόρμας για ελεύθερα συγγράμματα

Στην εργασία αυτή, θα πραγματοποιηθεί εκτενής αναζήτηση στη βιβλιογραφία και κατασκευή μιας διαδικτυακής πλατφόρμας (βάσης δεδομένων) που θα περιέχει όλα τα διαθέσιμα βιβλία ανοιχτής πρόσβασης (open access). Τα βιβλία θα ταξινομηθούν ανά κατηγορία και η βάση δεδομένων θα μπορεί να χρησιμοποιηθεί από μέλη της ακαδημαϊκής κοινότητας για αναζήτηση βιβλιογραφίας και διδακτικών εγχειριδίων, τόσο σαν κύριο όσο και σαν βοηθητικό διδακτικό υλικό, σε προπτυχιακά και μεταπτυχιακά μαθήματα.

Επιθυμητά προσόντα: γνώσεις εκπαιδευτικής έρευνας, γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος

7. Λογισμικό εύρεσης και μελέτης ανισορροπίας σύνδεσης

Οι μελέτες γενετικής συσχέτισης (GWAS), οι προσεγγίσεις fine-mapping και οι gene-based μέθοδοι στηρίζονται στην πληροφορία της ανισορροπίας σύνδεσης. Ωστόσο, τα υπάρχοντα εργαλεία για την εξερεύνηση προτύπων ανισορροπίας σύνδεσης παρουσιάζουν περιορισμούς ως προς το μέγεθος και τον τύπο των δεδομένων εισόδου, την διαθεσιμότητα, καθώς και ως προς την ποικιλομορφία των πληθυσμών στα πάνελ αναφοράς. Στο πλαίσιο της παρούσας διπλωματικής εργασίας θα επιχειρηθεί η ανάπτυξη ενός υπολογιστικού εργαλείου (γραμμής εντολών και διαδικτυακό) για την αποδοτική και ευέλικτη διερεύνηση προτύπων ανισορροπίας σύνδεσης από πολλαπλά πάνελ αναφοράς, με στόχο τη βέλτιστη αξιοποίηση της ανισορροπίας σύνδεσης σε αναλύσεις μελετών γενετικής συσχέτισης και συναφείς εφαρμογές.

Επιθυμητά προσόντα: Γνώσεις βιολογίας και βιοπληροφορικής, γνώσεις προγραμματισμού.

Επιβλέπων: Παντελής Μπάγκος