



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ

UNIVERSITY OF PIRAEUS
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGIES



Εργαστήριο Επιστήμης Δεδομένων Φοιτητικές υποτροφίες 2019-20 Πρόσκληση εκδήλωσης ενδιαφέροντος

Το Εργαστήριο Επιστήμης Δεδομένων της Σχολής Τεχνολογιών Πληροφορικής και Επικοινωνιών προσφέρει πέντε (5) υποτροφίες βραχείας διάρκειας που απευθύνονται σε προπτυχιακούς φοιτητές, μεταπτυχιακούς φοιτητές και Υποψήφιους Διδάκτορες του Παν/μίου Πειραιώς. Η κάθε υποτροφία έχει διάρκεια τριών (3) μηνών μέσα στο διάστημα Ιανουαρίου – Μαΐου 2020 και αντικείμενο την ανάπτυξη λογισμικού, όπως περιγράφεται παρακάτω στις επιμέρους εργασίες. Το ύψος της κάθε υποτροφίας κυμαίνεται μεταξύ 600-900 ευρώ (μικτά), ανάλογα με το εύρος της εργασίας, μέσω του Ειδικού Λογαριασμού Κονδυλίων Έρευνας (ΕΛΚΕ) του Παν/μίου. Η κάθε υποτροφία μπορεί να είναι ατομική ή σε ομάδα των 2 ατόμων.

Σε κάθε εργασία το παραδοτέο θα είναι μία τεχνική αναφορά και ένα GitHub repository (που θα περιλαμβάνει πηγαίο κώδικα, τεκμηρίωση κλπ.), το οποίο εφόσον κριθεί ότι πληροί τις σχετικές προδιαγραφές θα αναρτηθεί και στο υπό διαμόρφωση GitHub του Εργαστηρίου.

Για διευκρινίσεις σχετικά με τις εργασίες 1, 2, 3-5, οι ενδιαφερόμενοι μπορούν να απευθύνονται στους κκ. Καθηγητή Γιάννη Θεοδωρίδη (ytheod@unipi.gr), Επίκ. Καθηγητή Χρήστο Δουλκερίδη (cdoulk@unipi.gr), Επίκ. Καθηγητή Νίκο Πελέκη (npelekis@unipi.gr), αντίστοιχα.

Προθεσμία εκδήλωσης ενδιαφέροντος: Τετάρτη 15/1/2020, με αποστολή email στους ανωτέρω καθηγητές, στο οποίο θα δηλώνεται το θέμα (ή τα θέματα) για το οποίο υπάρχει ενδιαφέρον και θα επισυνάπτεται βιογραφικό σημείωμα και κατάλογος μαθημάτων (με την αντίστοιχη βαθμολογία) σχετικών με τη θεματολογία των εργασιών (Γλώσσες Προγραμματισμού, Βάσεις Δεδομένων, Αλγόριθμοι, Αναλυτική / Εξόρυξη Δεδομένων, Γεωγραφικά Συστήματα κλπ.).

1. Βιβλιοθήκη για προεπεξεργασία δεδομένων κίνησης

Στόχος της εργασίας είναι η ανάπτυξη μιας βιβλιοθήκης Python η οποία θα περιλαμβάνει μια σειρά μεθόδων και τεχνικών για την προ-επεξεργασία δεδομένων κίνησης (πρακτικά, τροχιών κινούμενων αντικειμένων) για σκοπούς αναλυτικής. Ενδεικτικές τεχνικές είναι: καθαρισμός πρωτογενών δεδομένων θέσης, επαναδειγματοληψία δεδομένων θέσης, τμηματοποίηση δεδομένων θέσης σε τροχιές, εξομάλυνση τροχιών, απλοποίηση τροχιών κλπ. Επίσης, υπολογισμός στατιστικών μέτρων πάνω σε διάφορα χαρακτηριστικά, όπως ενδεικτικά μήκος και διάρκεια τροχιάς, ταχύτητα, επιτάχυνση κ.α. Τέλος, υπολογισμός συναθροιστικής πληροφορίας, όπως μήτρα αφετηρίας-προορισμού, πλέγμα πυκνότητας κ.α. Σχετικές εργασίες (προφανώς, αρκετά πιο εκτεταμένες σε τεχνικό βάθος από την παραπάνω): `moving pandas` (<https://github.com/anitagraser/movingpandas>), `scikit-mobility` (<https://arxiv.org/pdf/1907.07062.pdf>).

2. Βιβλιοθήκη για οπτικοποίηση δεδομένων κίνησης

Στόχος της εργασίας είναι η ανάπτυξη ενός μηχανισμού που θα επιτρέπει την απεικόνιση σε χάρτη διαφόρων τύπων γεωχωρικών δεδομένων (σημεία, γραμμές, πολύγωνα, τροχιές κινούμενων αντικειμένων), με τρόπο εύκολο και παραμετροποιήσιμο ώστε να είναι εφικτή η επέκτασή του στο μέλλον. Ο μηχανισμός θα αποτελείται από μια βιβλιοθήκη λογισμικού που θα μπορεί (α) να τροφοδοτείται από διάφορες πηγές δεδομένων (π.χ. μια βάση δεδομένων PostGIS, αρχείο CSV, αρχείο Shapefile, κτλ.), και (β) να απεικονίζει τα γεωχωρικά δεδομένα σε ιστοσελίδα που θα περιέχει γεωγραφικό χάρτη, με διαφορετικά χρώματα, επισημειώσεις, και υποστηρίζοντας διάφορα φίλτρα επιλογής δεδομένων. Προτεινόμενη βιβλιοθήκη στην οποία θα στηριχθεί η υλοποίηση είναι η Leaflet (<https://leafletjs.com/>).

3. Εφαρμογή προβλέψεων για το Fantasy Challenge της Euroleague

Στόχος της εργασίας είναι η ανάπτυξη μίας εφαρμογής η οποία θα αναλύει τα δεδομένα του παιχνιδιού Fantasy Challenge της Euroleague, θα εφαρμόζει προβλεπτικές τεχνικές στα στατιστικά που επιτυγχάνουν οι παίκτες σε εβδομαδιαία βάση και θα προτείνει στους χρήστες της εφαρμογής τρόπους ενίσχυσης της ομάδας που οι ίδιοι διαχειρίζονται για την αναρρίχηση τους στη βαθμολογία του Fantasy Challenge. Ενδεικτικά, αναφέρεται η εφαρμογή τεχνικών πρόβλεψης της αξίας του κάθε παίκτη από χρονοσειρά των ιστορικών στατιστικών του. Περισσότερες πληροφορίες για το παιχνίδι στο URL: <https://fantasychallenge.euroleague.net/>.

4. Εφαρμογή χρονοπρογραμματισμού και πρόβλεψης κυκλοφοριακού φόρτου κρουαζιερόπλοιων

Στόχος της εργασίας είναι η ανάπτυξη μίας διαδικτυακής εφαρμογής η οποία θα επιτρέπει την καταγραφή του προγράμματος επίσκεψης των κρουαζιερόπλοιων και όλων των εκδρομικών σκαφών βάσει τοπικών λιμενικών κανονισμών, διεθνών κανονισμών ασφαλείας (όπως ο International Ship and Port Facility Security Code) και των περιορισμών που θέτουν οι τοπικές τουριστικές αρχές. Επιπρόσθετα θα επιτρέπει την απεικόνιση σε πραγματικό χρόνο των θέσεων των σκαφών και του θαλάσσιου κυκλοφοριακού φόρτου επιλεγμένων χώρων ελλιμενισμού ή αναμονής κρουαζιερόπλοιων (ή γενικότερα τουριστικών σκαφών), την πρόβλεψη αυτού σε προεπιλεγμένο χρονικό ορίζοντα αλλά και την πρόταση για προσαρμογή του προγράμματος επίσκεψης όταν προβλέπεται απόκλιση από τον καταγεγραμμένο προγραμματισμό που οδηγεί σε αυξημένο κυκλοφοριακό φόρτο.

5. Γεννήτορας συνθετικών δεδομένων κίνησης Hermoupolis

Το εργαλείο *Hermoupolis* είναι ένας γεννήτορας ρεαλιστικών, συνθετικών δεδομένων κίνησης που παράγει σημασιολογικά επισημειωμένες τροχιές κινούμενων αντικειμένων βάσει προδιαγεγραμμένων προτύπων κίνησης. Στόχος της εργασίας είναι η αντικατάσταση του παραδοσιακού γραφικού περιβάλλοντος χρήσης του εργαλείου από μία διαδικτυακή εφαρμογή που να απλοποιεί τη δημιουργία των προτύπων εισόδου με γραφικό τρόπο, δηλαδή με χρήση λειτουργιών πάνω σε χάρτες, καθώς και την αντικατάσταση του υπολογιστικού πλαισίου που τρέχει ο αλγόριθμος ώστε να επιτρέπονται προσομοιώσεις μεγάλης κλίμακας. Περισσότερες πληροφορίες για το εργαλείο *Hermoupolis* στο URL: http://infolab.cs.unipi.gr/?page_id=2251, ενώ σχετική εργασία αποτελεί και ο Minnesota Traffic Generator (MNTG), URL: <http://www.cs.ucr.edu/~amr/papers/sstd13.mntg.pdf>.

Το Εργαστήριο Επιστήμης Δεδομένων του Πανεπιστημίου Πειραιώς (Data Science Lab. – DataStories) δραστηριοποιείται ερευνητικά σε ένα πλήθος πεδίων της Επιστήμης Δεδομένων, όπως η διαχείριση μεγάλων δεδομένων, η στατιστική και η αναλυτική δεδομένων, η μηχανική μάθηση, η εξόρυξη γνώσης

από βάσεις δεδομένων, κειμένων και ήχων, η εξερεύνηση δεδομένων κίνησης και η ιδιωτικότητα δεδομένων. Το εργαστήριο έχει να επιδείξει μία σημαντική εμπειρία σε Ευρωπαϊκά ερευνητικά έργα, με πιο πρόσφατα τα H2020-ICT DATACRON (Big Data Analytics for Time Critical Mobility Forecasting, 2016-18, <http://datacron-project.eu>), H2020-ICT Track&Know (Big Data for Mobility Tracking Knowledge Extraction in Urban Areas, 2018-20, <https://trackandknowproject.eu>), και H2020-MSCA-RISE MASTER (Multiple Aspects Trajectory Management and Analysis, 2018-22, <http://www.master-project-h2020.eu>). Στην τρέχουσα σύνθεσή του, στο εργαστήριο δραστηριοποιούνται ερευνητικά 9 μέλη ΔΕΠ (4 διαφορετικών Τμημάτων) του Πανεπιστημίου Πειραιώς και απασχολείται αριθμός άνω των 10 εμπειρών και νέων ερευνητών. Περισσότερες πληροφορίες στο <http://www.datastories.org>.