

Θέμα: Μελέτη και υλοποίηση αλγορίθμων Μηχανικής Μάθησης σε δεδομένα ροής στο Apache Flink	
Επιβλέπων: Βασίλειος Ταμπακάς, καθηγητής	e-mail: tampakas@uop.gr τηλ:
Μέλη:	Ακαδημαϊκό Έτος: 2022-2023
<p>Στόχοι Σκοπός της παρούσας πτυχιακής εργασίας αποτελεί η μελέτη της τρέχουσας βιβλιογραφίας σχετικά με αλγόριθμους Μηχανικής Μάθησης καθώς και η υλοποίησή τους μέσω του Apache Flink. Ο φοιτητής θα υλοποιήσει μια ολοκληρωμένη εφαρμογή – η οποία θα εμπίπτει στα ενδιαφέροντα του φοιτητή - που θα εισάγει δεδομένα μέσω του Apache Kafka ή/και το Apache Flume, η δε αποθήκευση των επεξεργασμένων δεδομένων θα γίνεται στο HDFS, σε σχεσιακές βάσεις δεδομένων ή/και σε NoSQL datastores.</p>	
<p>Αντικείμενο Η επεξεργασία ροής (streaming) εμπλέκεται σε μεγάλο βαθμό με εφαρμογές σε μεγάλα δεδομένα (Big Data). Στον τομέα των μεγάλων δεδομένων, οι πιο σημαντικές και γνωστές μηχανές επεξεργασίας ροής είναι το Apache Flink, το Apache Storm και το Apache Spark.</p> <p>Στο Apache Spark και στο Apache Flink, το προγραμματιστικό μοντέλο έχει σχεδιαστεί πάνω σε μια αφαίρεση υψηλού επιπέδου. Αυτό επιτρέπει στον χρήστη να αναπτύσσει εφαρμογές πολύ πιο γρήγορα. Το Flink είναι ένα πλαίσιο επεξεργασίας δεδομένων τέταρτης γενιάς και είναι ένα από τα πιο γνωστά έργα Apache. Το Flink υποστηρίζει εγγενώς την επεξεργασία κατά δεσμίδες (batch processing) αλλά και την επεξεργασία ροής. Προωθεί τη συνεχή ροή και οι υπολογισμοί συμβάντων ενεργοποιούνται αμέσως μόλις ληφθεί το συμβάν.</p> <p>Αντικείμενο της πτυχιακής εργασίας είναι η εξοικείωση και χρήση με την τεχνολογία του Apache Flink. Θα αναπτυχθεί εφαρμογή, από άκρο σε άκρο που αντιδρά σε δεδομένα σε πραγματικό χρόνο, η οποία με χρήση του Flink και των αλγορίθμων Μηχανικής Μάθησης θα προσφέρει in time αναλυτικές υπηρεσίες (in time big data analytics).</p> <p>Η γλώσσα υλοποίησης της εφαρμογής μπορεί να είναι η Java, η Scala ή η Python ανάλογα με την εμπειρία του/των φοιτητών.</p>	
<p>Η εργασία περιλαμβάνει: (π.χ. σχεδιασμό συστήματος, ανάπτυξη συστήματος, διερεύνηση βιβλιογραφίας κ.λ.π) [✓] Σχεδιασμό και ανάπτυξη συστήματος [✓] Συγκριτική μελέτη και πλαίσιο αξιολόγησης [✓] Ανάλυση και σχεδιασμό μοντέλου</p>	
<p>Σχετιζόμενα Μαθήματα Λειτουργικά Συστήματα, Εισαγωγή στα Καταναεμημένα συστήματα, Συστήματα Διαχείρισης Δεδομένων</p>	

Απαιτήσεις παρουσίας φοιτητή: ΟΧΙ

Επιπλέον Υποστήριξη πτυχιακής: Μπάλλας Ιωάννης, Υποψήφιος διδάκτωρ, Ερευνητικό εργαστήριο DiSyD