

ΠΕΡΙΓΡΑΜΜΑ ΜΑΘΗΜΑΤΟΣ

(1) ΓΕΝΙΚΑ

ΣΧΟΛΗ	Μηχανικών		
ΤΜΗΜΑ	Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	Προπτυχιακό		
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	0811.9.017.0	ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ	9 ^ο
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	Μεγάλα Δεδομένα		
ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ	ΕΒΔΟΜΑΔΙΑΙΕΣ ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ	ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ	
Διαλέξεις και Ασκήσεις Πράξης	4	3	
Εργαστηριακές Ασκήσεις	1	1	
ΣΥΝΟΛΟ	5	4	
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ	Εμβάθυνσης / Εμπέδωσης γνώσεων ειδικότητας		
ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ	Δομημένος Προγραμματισμός, Αντικειμενοστραφής Προγραμματισμός, Βάσεις Δεδομένων, Συστήματα Αξιολόγησης Διαδικτυακών Εφαρμογών		
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ	Ελληνική		
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	Ναι (στα Αγγλικά & Γερμανικά)		
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)	https://eclass.hmu.gr/courses/ECE194/		

(2) ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Μαθησιακά Αποτελέσματα
<p>Ο βασικός στόχος του μαθήματος είναι η κατανόηση και η εκμάθηση τεχνικών επεξεργασίας μεγάλων δεδομένων. Στο μάθημα γίνεται ανάλυση των παραγόντων που οδηγούν σε πολύ μεγάλους όγκους δεδομένων, ροής δεδομένων και της δομής τους. Γίνεται παρουσίαση των τεχνικών, των εργαλείων και του λογισμικού που χρησιμοποιείται σήμερα για την αξιοποίηση των μεγάλων δεδομένων. Αναλύεται το οικοσύστημα του Hadoop γίνεται παρουσίαση του τρόπου εγκατάστασης και χρήσης του Hadoop File System. Παρουσιάζουμε το Hive, το μοντέλο MapReduce με παραδείγματα και τέλος το Apache Spark. Μια σειρά από άλλα χρήσιμα εργαλεία υποδομής, όπως NoSQL βάσεις δεδομένων (HBase) παρουσιάζονται επίσης, καθώς και υπηρεσίες και εφαρμογές που χρησιμοποιούνται για την γεφύρωση των κλασικών πληροφοριακών συστημάτων με τις τεχνολογίες μεγάλων δεδομένων.</p> <p>Με την επιτυχή ολοκλήρωση του μαθήματος ο φοιτητής / τρια θα είναι σε θέση να:</p> <ul style="list-style-type: none"> ▪ Κατανοεί τον τρόπο επεξεργασίας των μεγάλων δεδομένων ▪ Εγκαθιστά Hadoop File System σε απλό υπολογιστή ή σε cluster υπολογιστών ▪ Εφαρμόζει τεχνικές MapReduce για εξαγωγή αποτελεσμάτων αναζήτησης. ▪ Σχεδιάζει εφαρμογές που χρησιμοποιούν NoSQL βάσεις δεδομένων ▪ Εγκαθιστά Zookeeper και να μπορεί να το αξιοποιήσει στο HDFS ▪ Γνωρίζει το οικοσύστημα Apache Spark και να μπορεί να χρησιμοποιεί Spark SQL.
Γενικές Ικανότητες
<ul style="list-style-type: none"> • Αυτόνομη εργασία • Ομαδική εργασία • Προαγωγή της δημιουργικότητας

(3) ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ

Ενότητες Θεωρητικών Διαλέξεων

- Εισαγωγή στα μεγάλα δεδομένα
- Hadoop File System
- Αρχιτεκτονική, τοπολογία και εναλλακτικές εφαρμογές
- Zookeeper
- Key/Value Stores
- Οριζόντιο και κάθετο scaling
- Sharding
- NoSQL Βάσεις Δεδομένων (HBase, Cassandra)
- MapReduce, Hadoop
- Hive
- Pig
- Kafka
- Amazon S3

Εργαστηριακές Ασκήσεις

Στο εργαστηριακό μέρος του μαθήματος οι φοιτητές έχουν τη δυνατότητα της πρακτικής εφαρμογής των εννοιών της θεωρίας με τη χρήση ασκήσεων που καλύπτουν εκτενώς την ύλη και καλλιεργούν ορθές δεξιότητες για την εγκατάσταση, χρήση και ανάπτυξη εφαρμογών μεγάλων δεδομένων με σύγχρονες τεχνικές.

(4) ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ.	Στην τάξη	
ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ	Στη Διδασκαλία Στην Επικοινωνία με τους φοιτητές Classweb, e-mail	
ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ	Δραστηριότητα	Φόρτος Εργασίας Εξαμήνου
	Διαλέξεις	45
	Μικρές ατομικές εργασίες εξάσκησης	30
	Αυτοτελής Μελέτη	40
	Γραπτή Εξέταση	5
	Σύνολο Μαθήματος	120
ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ	Γλώσσα Αξιολόγησης: Ελληνική Θεωρία: Τελική γραπτή εξέταση στο σύνολο της ύλης (100%). Η εξέταση περιλαμβάνει ερωτήματα θεωρίας (από 3 έως 5) και ασκήσεις πράξεις (από 1 έως 2). Εργαστήριο: Ο τελικός βαθμός συναθροίζεται από συγγραφή εργαστηριακών εργασιών (30%), εκπόνηση project (50%) και τελικής εξέτασης (20%) Τα κριτήρια αξιολόγησης ανακοινώνονται στους φοιτητές κατά την έναρξη του εξαμήνου και βρίσκονται αναρτημένα στην ιστοσελίδα του μαθήματος στο eClass.	

(5) ΣΥΝΙΣΤΩΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

- Προτεινόμενη Βιβλιογραφία:

- Raman Jhajj (2016): Apache Hadoop Cookbook. Java Code Geeks
- Steve Hoffman (2015): Apache Flume: Distributed Log Collection for Hadoop Second Edition, Packt Publishing

- Martin Mois (2017): Amazon S3 Tutorial, Hot Recipes for the Amazon S3 Platform. Java Code Geeks
- Balaswamy Vaddeman (2016): Beginning Apache Pig, Apress

-Συναφή επιστημονικά περιοδικά:

- Journal of Big Data, Springer
- Journal of Big Data Research, Elsevier
- International Journal of Big Data Intelligence, Inderscience
- Frontiers in Big Data