

ΠΕΡΙΓΡΑΜΜΑ ΜΑΘΗΜΑΤΟΣ

1. ΓΕΝΙΚΑ

ΣΧΟΛΗ	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ		
ΤΜΗΜΑ/ΣΥΜΜΕΤΕΧΟΝΤΑ ΤΜΗΜΑΤΑ*	ΜΑΘΗΜΑΤΙΚΩΝ		
ΣΥΜΜΕΤΕΧΟΝΤΑ ΙΔΡΥΜΑΤΑ**	Βάσεις Δεδομένων και Εξόρυξη Δεδομένων		
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ: ΤΙΤΛΟΣ Π.Μ.Σ.	ΥΠΟΛΟΓΙΣΤΙΚΗ ΚΑΙ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΤΙΚΗ ΣΤΗΝ ΕΠΙΣΤΗΜΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ (MCDA)		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΜΕΤΑΠΤΥΧΙΑΚΟ		
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	MCDA203	ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ	B
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΕΞΟΡΥΞΗ ΔΕΔΟΜΕΝΩΝ		
ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ <i>σε περίπτωση που οι πιστωτικές μονάδες απονέμονται σε διακριτά μέρη του μαθήματος π.χ. Διαλέξεις, Εργαστηριακές Ασκήσεις κ.λπ. Αν οι πιστωτικές μονάδες απονέμονται ενιαία για το σύνολο του μαθήματος αναγράψτε τις εβδομαδιαίες ώρες διδασκαλίας και το σύνολο των πιστωτικών μονάδων</i>	ΕΒΔΟΜΑΔΙΑΙΕΣ ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ	ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ	
Διαλέξεις	2	7.5	
Εργαστήριο	1		
<i>Προσθέστε σειρές αν χρειαστεί. Η οργάνωση διδασκαλίας και οι διδακτικές μέθοδοι που χρησιμοποιούνται περιγράφονται αναλυτικά στο 4.</i>			
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ <i>Υποβάθρου, Γενικών Γνώσεων, Επιστημονικής Περιοχής, Ανάπτυξης Δεξιοτήτων</i>	Υποβάθρου		
ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ:	Όχι		
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:	Ελληνική		
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	Ναι		
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)	https://eclass.math.upatras.gr/courses/MATHDEP233/		

2. ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

<p>Μαθησιακά Αποτελέσματα</p> <p>Περιγράφονται τα μαθησιακά αποτελέσματα του μαθήματος οι συγκεκριμένες γνώσεις, δεξιότητες και ικανότητες καταλλήλου επιπέδου που θα αποκτήσουν οι φοιτητές μετά την επιτυχή ολοκλήρωση του μαθήματος.</p> <p>Συμβουλευτείτε το Παράρτημα Α (ξεχωριστό αρχείο στο e-mail)</p> <ul style="list-style-type: none"> • Περιγραφή του Επιπέδου των Μαθησιακών Αποτελεσμάτων για κάθε ένα κύκλο σπουδών σύμφωνα με Πλαίσιο Προσόντων του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης • Περιγραφικοί Δείκτες Επιπέδων 6, 7 & 8 του Ευρωπαϊκού Πλαισίου Προσόντων Διά Βίου Μάθησης και Παράρτημα Β • Περιληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων
<p>Με το μάθημα αυτό ο φοιτητής θα μπορεί να χρησιμοποιεί υπολογιστικές μεθόδους για την επίλυση προβλημάτων της Επιστήμης Δεδομένων. Στο τέλος αυτού του μαθήματος ο φοιτητής θα έχει περαιτέρω αναπτύξει τις ακόλουθες δεξιότητες:</p> <ul style="list-style-type: none"> • προετοιμασία, καθαρισμός και οπτικοποίηση των δεδομένων συνδυάζοντας SQL και R, • κατανόηση των μεθόδων επιτηρούμενης, ημι-επιτηρούμενης και μη επιτηρούμενης μάθησης, • ικανότητα διάκρισης των διαφορών μεταξύ των μεθόδων προκειμένου να μπορεί να επιλέξει την καταλληλότερη για το πρόβλημα που καλείται να επιλύσει, • ικανότητα εφαρμογής αυτών των μεθόδων στη λύση προβλημάτων με χρήση της γλώσσας R.

Μετά την επιτυχή παρακολούθηση του μαθήματος, ο φοιτητής θα είναι σε θέση να δίνει λύσεις σε πραγματικά προβλήματα χρησιμοποιώντας μία κατάλληλη μέθοδο εξόρυξης δεδομένων.

Γενικές Ικανότητες

Λαμβάνοντας υπόψη τις γενικές ικανότητες που πρέπει να έχει αποκτήσει ο πτυχιούχος (όπως αυτές αναγράφονται στο Παράρτημα Διπλώματος και παρατίθενται ακολούθως) σε ποια / ποιες από αυτές αποσκοπεί το μάθημα;

Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών

Προσαρμογή σε νέες καταστάσεις

Λήψη αποφάσεων

Αυτόνομη εργασία

Ομαδική εργασία

Εργασία σε διεθνές περιβάλλον

Εργασία σε διεπιστημονικό περιβάλλον

Παράγωγή νέων ερευνητικών ιδεών

Σχεδιασμός και διαχείριση έργων

Σεβασμός στη διαφορετικότητα και στην πολυπολιτισμικότητα

Σεβασμός στο φυσικό περιβάλλον

Επίδειξη κοινωνικής, επαγγελματικής και ηθικής υπευθυνότητας και ευαισθησίας σε θέματα φύλου

Άσκηση κριτικής και αυτοκριτικής

Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης

- Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών.
- Προσαρμογή σε νέες καταστάσεις.
- Λήψη αποφάσεων.
- Αυτόνομη εργασία.
- Ομαδική εργασία.
- Εργασία σε διεπιστημονικό περιβάλλον.
- Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης.

3. ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ

Μέρος Α: Θεωρία

(i) Εισαγωγή στις Βάσεις Δεδομένων, SQL. (ii) Προετοιμασία Δεδομένων. Καθαρισμός των δεδομένων (data cleaning). Συμπλήρωση ελλειπών τιμών. (iii) Επιβλεπόμενη μάθηση: Δέντρα απόφασης, Κοντινότεροι γείτονες, μπεϋζιανοί αλγόριθμοι, σύγκριση μεθόδων ταξινόμησης, ομάδες ταξινομητών. (iv) Παλινδρόμηση: Πολλαπλή γραμμική παλινδρόμηση, Δέντρα προβλέψεων, Νευρωνικά Δίκτυα. (v) Επιλογή ανεξάρτητων μεταβλητών: Μέθοδοι φίλτρου και περιτυλίγματος, Μέθοδος Ανάλυσης Κύριων Συνιστωσών. (vi) Μη-επιβλεπόμενη μάθηση: Συσταδοποίηση (Clustering), Μοντέλα Ιεραρχικής Συσταδοποίησης και Πυκνότητας. (vii) Κανόνες συσχέτισης, Αραιοί πίνακες. (viii) Μεγάλα Δεδομένα (Big Data): Hadoop και MapReduce.

Μέρος Β: Εργαστήριο

Η γλώσσα R για την Επιστήμη των Δεδομένων, Queries σε πολλαπλούς πίνακες με την εντολή JOIN, Τελεστές. Subqueries. Σύνδεση με R, Χρήση πακέτων της R: sqldf, lattice, ggplot2, dplyr, party, C50, Rattle, mlr, randomForest, rpart, caret, factoextra, cluster, fpc, arules, arulesViz, RHadoop.

4. ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

<p>ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.</p>	<p>Πρόσωπο με πρόσωπο</p>	
<p>ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ Χρήση Τ.Π.Ε. στη Διδασκαλία, στην Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές</p>	<ul style="list-style-type: none"> • Χρήση Τ.Π.Ε. <ul style="list-style-type: none"> ✓ στη διδασκαλία με ηλεκτρονικές διαφάνειες, ✓ στην εκμάθηση γλώσσας R. • Υποστήριξη της μαθησιακής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας eClass. 	
<p>ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι διδασκαλίας. Διαλέξεις, Σεμινάρια, Εργαστηριακή Άσκηση, Άσκηση Πεδίου, Μελέτη & ανάλυση βιβλιογραφίας, Φροντιστήριο, Πρακτική (Τοποθέτηση), Κλινική Άσκηση, Καλλιτεχνικό Εργαστήριο, Διαδραστική διδασκαλία, Εκπαιδευτικές επισκέψεις, Εκπόνηση μελέτης (project), Συγγραφή εργασίας / εργασιών, Καλλιτεχνική δημιουργία, κ.λπ. Αναγράφονται οι ώρες μελέτης του φοιτητή για κάθε μαθησιακή δραστηριότητα καθώς και οι ώρες μη καθοδηγούμενης μελέτης ώστε ο συνολικός φόρτος εργασίας σε επίπεδο εξαμήνου να αντιστοιχεί στα standards του ECTS</p>	Δραστηριότητα	Φόρτος Εργασίας Εξαμήνου
	<p>Διαλέξεις</p>	26
	<p>Εργαστήριο</p>	13
	<p>Μελέτη (μη καθοδηγούμενη)</p>	100
	<p>Εκπόνηση εργασιών</p>	45
<p>Εξέταση</p>	2.5	
<p>Εξέταση εργαστηρίου</p>	1	
<p>Σύνολο Μαθήματος (25 ώρες φόρτου εργασίας ανά πιστωτική μονάδα)</p>	187.5	
<p>ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ Περιγραφή της διαδικασίας αξιολόγησης Γλώσσα Αξιολόγησης, Μέθοδοι αξιολόγησης, Διαμορφωτική ή Συμπερασματική, Δοκιμασία Πολλαπλής Επιλογής, Ερωτήσεις Σύντομης Απάντησης, Ερωτήσεις Ανάπτυξης Δοκιμίων, Επίλυση Προβλημάτων, Γραπτή Εργασία, Έκθεση / Αναφορά, Προφορική Εξέταση, Δημόσια Παρουσίαση, Εργαστηριακή Εργασία, Κλινική Εξέταση Ασθενούς, Καλλιτεχνική Ερμηνεία, Άλλη / Άλλες Αναφέρονται ρητά προσδιορισμένα κριτήρια αξιολόγησης και εάν και που είναι προσβάσιμα από τους φοιτητές;</p>	<ul style="list-style-type: none"> • Γραπτή εξέταση (50%) • Εξέταση στο εργαστήριο (25%) • Εργασίες (25%) <p>Μικρότερος προβιβάζσιμος βαθμός: 5 Μέγιστος προβιβάζσιμος βαθμός: 10</p>	

5. ΣΥΝΙΣΤΩΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

<ul style="list-style-type: none"> • Beard, B. (2016). <i>Beginning SQL Server R Services: Analytics for Data Scientists</i>. Apress. • Torgo, L. (2016). <i>Data Mining With R: Learning With Case Studies</i>. CRC press. • Wickham, H., and Golemund, G. (2016). <i>R for Data Science: Import, Tidy, Transform, Visualize, and Model Data</i>. O'Reilly Media, Inc. • Zaki, M.J. and Wagber, M. Jr. (2017). <i>Εξόρυξη και Ανάλυση Δεδομένων: Βασικές Έννοιες και Αλγόριθμοι</i>. Εκδόσεις Κλειδάριθμος ΕΠΕ. • Βερούκιος, Β., Καγκλής, Β., και Σταυρόπουλος, Η. (2015). <i>Η Επιστήμη των Δεδομένων Μέσα από τη Γλώσσα R</i>. [ηλεκτρ. βιβλ.] Σύνδεσμος Ελληνικών Ακαδημαϊκών Βιβλιοθηκών. Διαθέσιμο στο http://hdl.handle.net/11419/2965
