

## ΠΕΡΙΓΡΑΜΜΑ ΜΑΘΗΜΑΤΟΣ

### (1) ΓΕΝΙΚΑ

<b>ΣΧΟΛΗ</b>	ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ		
<b>ΤΜΗΜΑ</b>	ΜΑΘΗΜΑΤΙΚΩΝ		
<b>ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ</b>	ΠΡΟΠΤΥΧΙΑΚΟ		
<b>ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ</b>	82405_8ΕΠ01	ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ	8 <sup>ο</sup>
<b>ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ</b>	ΕΞΟΡΥΞΗ ΚΑΙ ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΜΕΓΑΛΟΥ ΟΓΚΟΥ		
<b>ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ</b> <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>		<b>ΕΒΔΟΜΑΔΙΑΙΕΣ ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ</b>	<b>ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ</b>
Διαλέξεις		3	5
Προσθέστε σειρές αν χρειαστεί. Η οργάνωση διδασκαλίας και οι διδακτικές μέθοδοι που χρησιμοποιούνται περιγράφονται αναλυτικά στο (δ).		3	5
<b>ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ</b> <i>γενικού υποβάθρου, ειδικού υποβάθρου, ειδίκευσης γενικών γνώσεων, ανάπτυξης δεξιοτήτων</i>	Μάθημα επιλογής, ανάπτυξης δεξιοτήτων που ανήκει στην επιστημονική περιοχή «Υπολογιστικά & Εφαρμοσμένα Μαθηματικά»		
<b>ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ:</b>	Όχι		
<b>ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ:</b>	Ελληνική		
<b>ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS</b>	Ναι (στην αγγλική γλώσσα, για φοιτητές Erasmus)		
<b>ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)</b>	<a href="http://math.uth.gr/?page_id=877">http://math.uth.gr/?page_id=877</a>		

### (2) ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

<p><b>Μαθησιακά Αποτελέσματα</b>  <i>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.</i></p> <p><i>Συμβουλευτείτε το Παράρτημα Α</i></p> <ul style="list-style-type: none"> <li>• Περιγραφή του Επιπέδου των Μαθησιακών Αποτελεσμάτων για κάθε ένα κύκλο σπουδών σύμφωνα με το Πλαίσιο Προσόντων του Ευρωπαϊκού Χώρου Ανώτατης Εκπαίδευσης</li> <li>• Περιγραφικοί Δείκτες Επιπέδων 6, 7 &amp; 8 του Ευρωπαϊκού Πλαισίου Προσόντων Διά Βίου Μάθησης και το Παράρτημα Β</li> <li>• Περιληπτικός Οδηγός συγγραφής Μαθησιακών Αποτελεσμάτων</li> </ul>
<p>Στο μάθημα γίνεται μια εισαγωγή στην εξόρυξη πληροφορίας και την ανακάλυψη γνώσης από δεδομένα που χαρακτηρίζονται από μεγάλο πλήθος δειγμάτων, υψηλή διάσταση και την ταχύ ροή νέων παρατηρήσεων, ενσωματώνοντας έννοιες από τη Μηχανική Μάθηση, τη Στατιστική και τη Γραμμική Άλγεβρα. Οι βασικές μέθοδοι ομαδοποίησης, κατηγοριοποίησης και πρόβλεψης αναλύονται, καθώς και τα απαραίτητα πρακτικά εργαλεία για την εφαρμογή τους.</p> <p>Τα κύρια μέρη του μαθήματος περιλαμβάνουν τη διερευνητική ανάλυση δεδομένων, την επιβλεπόμενη και μη επιβλεπόμενη μάθηση, την ανάλυση ροών δεδομένων και εφαρμογές αυτών σε πραγματικά προβλήματα της βιομηχανίας. Το μάθημα θέτει τα βασικά θεμέλια αυτών των περιοχών, αλλά καλύπτει επίσης θέματα αιχμής, όπως είναι οι μέθοδοι πυρήνα, η ανάλυση μεγάλων διαστάσεων δεδομένων μέσω προβολών και η ανάλυση γράφων και δικτύων. Επίσης παρουσιάζονται βασικές μέθοδοι για την καταγραφή και αποθήκευση δεδομένων μεγάλου όγκου, μαζί με τις βασικές εφαρμογές τους στην εξόρυξη κοινωνικών δικτύων και στην εξόρυξη κειμένου. Το μάθημα καταλήγει με μια εισαγωγή στην οπτικοποίηση δεδομένων μεγάλου όγκου.</p> <p>Οι ασκήσεις του μαθήματος επικεντρώνονται στη χρήση εργαλείων ανοιχτού λογισμικού για τη διαχείριση και ανάλυση δεδομένων μεγάλου όγκου που προέρχονται από πραγματικά προβλήματα.</p>

Με την επιτυχή παρακολούθηση και ολοκλήρωση του μαθήματος η φοιτήτρια/ο φοιτητής θα είναι σε θέση:

- Να εκτελεί μια διερευνητική αριθμητική και γραφική ανάλυση ενός συνόλου δεδομένων μεγάλου όγκου.
- Να ταξινομεί μελλοντικές πολυμεταβλητές παρατηρήσεις σε έναν από τους πολλούς γνωστούς πληθυσμούς.
- Να εφαρμόζει διαφορετικούς τύπους αλγορίθμων εξόρυξης δεδομένων σύμφωνα με τους διαφορετικούς τύπους προβλημάτων.
- Να κατανοεί τους υπολογιστικούς περιορισμούς των διαφόρων αλγορίθμων, όταν εφαρμόζονται σε δεδομένα μεγάλου όγκου.
- Να εφαρμόζει λύσεις σε προβλήματα πραγματικών εφαρμογών που παράγουν δεδομένα μεγάλου όγκου.
- Να χρησιμοποιεί εργαλεία ανοιχτού λογισμικού (R-project) για τους παραπάνω στόχους.

#### 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

.....

Άλλες...

.....

Με την επιτυχή παρακολούθηση και ολοκλήρωση του μαθήματος η φοιτήτρια/ο φοιτητής θα έχει αποκτήσει τις ακόλουθες ικανότητες:

- Αυτόνομη εργασία
- Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης
- Αναλυτική και συνθετική σκέψη
- Κριτική σκέψη
- Επίλυση προβλημάτων
- Παραγωγή νέων ερευνητικών ιδεών
- Αναζήτηση, ανάλυση και σύνθεση πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών

### (3) ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ

- Γραφική απεικόνιση και περίληψη δεδομένων μεγάλου όγκου.
- Η κανονική κατανομή. Γραμμικοί συνδυασμοί τυχαίων μεταβλητών. Γραμμική ανάλυση διακρίσεων. Συζήτηση για την εκτίμηση του ποσοστού σφάλματος.
- Αλγόριθμοι επιβλεπόμενης μάθησης. naive Bayes, kNN, Support Vector Machine (SVM), συνδυαστικές μέθοδοι.
- Ομαδοποίηση. Ταξινόμηση τεχνικών ομαδοποίησης. Μέθοδοι ομαδοποίησης με βάση το κέντρο βάρους (k-means), ιεραρχική ομαδοποίηση (συγχωνευτική και διχαστική), ομαδοποίηση βασισμένη στην πυκνότητα (DBSCAN).
- Εισαγωγή στην ανάλυση πρωτευουσών συνιστωσών και εφαρμογές.
- Μείωση διάστασης σε δεδομένα μεγάλου όγκου (Τυχαίες Προβολές, Παράλληλες μέθοδοι). Μείωση διάστασης για την ομαδοποίηση.
- Εξόρυξη προτύπων για την αναγνώριση κοινωνικής συμπεριφοράς στο διαδίκτυο.
- Εξόρυξη προτύπων σε δεδομένα κειμένου μεγάλου όγκου. Κατηγοριοποίηση αρχείων κειμένου και γρήγορη αναζήτηση.
- Εφαρμογές στη Βιοϊατρική, γονιδιακές εκφράσεις πληθυσμών, εξόρυξη σε ακολουθίες.

#### (4) ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ - ΑΞΙΟΛΟΓΗΣΗ

<b>ΤΡΟΠΟΣ ΠΑΡΑΔΟΣΗΣ</b> <i>Πρόσωπο με πρόσωπο, Εξ αποστάσεως εκπαίδευση κ.λπ.</i>	Πρόσωπο με πρόσωπο, στο αμφιθέατρο.												
<b>ΧΡΗΣΗ ΤΕΧΝΟΛΟΓΙΩΝ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ</b> <i>Χρήση Τ.Π.Ε. στη Διδασκαλία, στην Εργαστηριακή Εκπαίδευση, στην Επικοινωνία με τους φοιτητές</i>	Υποστήριξη εκπαιδευτικής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας e-class  Χρήση Τ.Π.Ε. στην επικοινωνία με τους φοιτητές (e-mail, ανακοινώσεις μέσω της ηλεκτρονικής πλατφόρμας e-class)  Υποστήριξη Μαθησιακής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας e-class												
<b>ΟΡΓΑΝΩΣΗ ΔΙΔΑΣΚΑΛΙΑΣ</b> <i>Περιγράφονται αναλυτικά ο τρόπος και μέθοδοι διδασκαλίας. Διαλέξεις, Σεμινάρια, Εργαστηριακή Άσκηση, Άσκηση Πεδίου, Μελέτη &amp; 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 σύμφωνα με τις αρχές του ECTS</i>	<table border="1"><thead><tr><th><b>Δραστηριότητα</b></th><th><b>Φόρτος Εργασίας Εξαμήνου</b></th></tr></thead><tbody><tr><td>Διαλέξεις</td><td>39</td></tr><tr><td>Μελέτη θεωρίας</td><td>33</td></tr><tr><td>Μελέτη, προετοιμασία και συγγραφή εργασιών</td><td>23</td></tr><tr><td>Μελέτη για τελική εξέταση</td><td>30</td></tr><tr><td>Σύνολο Μαθήματος (25 ώρες φόρτου εργασίας ανά πιστωτική μονάδα)</td><td>125</td></tr></tbody></table>	<b>Δραστηριότητα</b>	<b>Φόρτος Εργασίας Εξαμήνου</b>	Διαλέξεις	39	Μελέτη θεωρίας	33	Μελέτη, προετοιμασία και συγγραφή εργασιών	23	Μελέτη για τελική εξέταση	30	Σύνολο Μαθήματος (25 ώρες φόρτου εργασίας ανά πιστωτική μονάδα)	125
	<b>Δραστηριότητα</b>	<b>Φόρτος Εργασίας Εξαμήνου</b>											
	Διαλέξεις	39											
	Μελέτη θεωρίας	33											
	Μελέτη, προετοιμασία και συγγραφή εργασιών	23											
	Μελέτη για τελική εξέταση	30											
Σύνολο Μαθήματος (25 ώρες φόρτου εργασίας ανά πιστωτική μονάδα)	125												
<b>ΑΞΙΟΛΟΓΗΣΗ ΦΟΙΤΗΤΩΝ</b> <i>Περιγραφή της διαδικασίας αξιολόγησης</i>	1. Γραπτή τελική εξέταση στην ελληνική γλώσσα με τη μορφή: <ul style="list-style-type: none"><li>Ερωτήσεων ανάπτυξης.</li><li>Ερωτήσεων ανοικτού τύπου/Σύντομης απάντησης/πολλαπλής επιλογής με πλήρη τεκμηρίωση των απαντήσεων.</li></ul> 2. Ατομικές εργασίες, η βαθμολογία των οποίων συυπολογίζεται στον τελικό βαθμό. 3. Προφορικές εξετάσεις (όταν προβλέπεται). 4. Ο τρόπος και τα κριτήρια αξιολόγησης είναι προσβάσιμα από τις φοιτήτριες και τους φοιτητές μέσω της πλατφόρμας e-class.												
<i>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</i>													
<i>Αναφέρονται ρητά προσδιορισμένα κριτήρια αξιολόγησης και εάν και που είναι προσβάσιμα από τους φοιτητές.</i>													

#### (5) ΣΥΝΙΣΤΩΜΕΝΗ-ΒΙΒΛΙΟΓΡΑΦΙΑ

<ol style="list-style-type: none"><li>1. M.H. Dunham, Data Mining, Εισαγωγικά και Προηγμένα θέματα Εξόρυξης Γνώσης από Δεδομένα, Εκδ. ΕΚΔΟΣΕΙΣ ΝΕΩΝ ΤΕΧΝΟΛΟΓΙΩΝ ΙΔΙΩΤΙΚΗ ΚΕΦΑΛΑΙΟΥΧΙΚΗ ΕΤΑΙΡΕΙΑ, 1<sup>η</sup> εκδ., 2004. Κωδικός βιβλίου στον Εύδοξο: 395</li><li>2. Αλ. Νανόπουλος – Γ. Μανωλόπουλος, Εισαγωγή στην Εξόρυξη Δεδομένων και τις Αποθήκες Δεδομένων. Κωδικός βιβλίου στον Εύδοξο: 3079</li><li>3. M.J. Zaki, M.JR. Wagner, Εξόρυξη και Ανάλυση Δεδομένων: βασικές έννοιες και αλγόριθμοι. Κωδικός βιβλίου στον Εύδοξο: 68386089</li><li>4. Tan Pang – Ning, Steinbach Michael, Kumar Vipin, Εισαγωγή στην εξόρυξη δεδομένων. Κωδικός βιβλίου στον Εύδοξο: 18549105</li><li>5. Hastie, T., Tibshirani, R. and Friedman, J., <i>The Elements of Statistical Learning: Data Mining, Inference, and Prediction</i>, Springer, 2<sup>nd</sup> Edition, 2011.</li><li>6. J. Gareth, D. Witten, Tr. Hastie and R. Tibshirani, <i>An Introduction to Statistical Learning with Applications in R</i>, Springer, 2013.</li></ol>
---