

5.6 Μαθήματα ΣΤ' εξαμήνου

5.6.1 Εξόρυξη Γνώσης και Μηχανική Μάθηση

(α) ΓΕΝΙΚΑ

ΣΧΟΛΗ	ΟΙΚΟΝΟΜΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ		
ΤΜΗΜΑ	ΨΗΦΙΑΚΩΝ ΣΥΣΤΗΜΑΤΩΝ		
ΕΠΙΠΕΔΟ ΣΠΟΥΔΩΝ	ΠΡΟΠΤΥΧΙΑΚΟ		
ΚΩΔΙΚΟΣ ΜΑΘΗΜΑΤΟΣ	ΨΣ026		
ΕΞΑΜΗΝΟ ΣΠΟΥΔΩΝ	ΣΤ		
ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ	ΕΞΟΡΥΞΗ ΓΝΩΣΗΣ ΚΑΙ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ		
ΑΥΤΟΤΕΛΕΙΣ ΔΙΔΑΚΤΙΚΕΣ ΔΡΑΣΤΗΡΙΟΤΗΤΕΣ	ΕΒΔΟΜΑΔΙΑΙΕΣ ΩΡΕΣ ΔΙΔΑΣΚΑΛΙΑΣ	ΠΙΣΤΩΤΙΚΕΣ ΜΟΝΑΔΕΣ	
Διαλέξεις	3		
Φροντιστηριακή διδασκαλία	1		
Σύνολο	4	6	
ΤΥΠΟΣ ΜΑΘΗΜΑΤΟΣ	υποχρεωτικό, επιστημονικής περιοχής (ειδικού υποβάθρου), μάθημα με φροντιστήριο		
ΠΡΟΑΠΑΙΤΟΥΜΕΝΑ ΜΑΘΗΜΑΤΑ			
ΓΛΩΣΣΑ ΔΙΔΑΣΚΑΛΙΑΣ και ΕΞΕΤΑΣΕΩΝ	Ελληνική		
ΤΟ ΜΑΘΗΜΑ ΠΡΟΣΦΕΡΕΤΑΙ ΣΕ ΦΟΙΤΗΤΕΣ ERASMUS	Ναι (στην Αγγλική)		
ΗΛΕΚΤΡΟΝΙΚΗ ΣΕΛΙΔΑ ΜΑΘΗΜΑΤΟΣ (URL)	https://eclass.uop.gr/modules/auth/opencourses.php?fc=294		

(β) ΜΑΘΗΣΙΑΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

■ Μαθησιακά Αποτελέσματα

Σκοπός του μαθήματος είναι η γνωριμία των φοιτητών και των φοιτητριών με έννοιες που αφορούν την εξόρυξη γνώσης από δεδομένων και τη μηχανική μάθηση. Πιο συγκεκριμένα, στόχοι του μαθήματος αποτελούν το να βοηθήσει τους φοιτητές και τις φοιτήτριες:

- Να γνωρίσουν τις εφαρμογές Εξόρυξης Γνώσης και Μηχανικής Μάθησης καθώς και τα στάδια της Εξόρυξης Γνώσης από Δεδομένα (Knowledge Discovery in Databases)
- Να εφαρμόζουν τις κατάλληλες τεχνικές προ-επεξεργασίας δεδομένων με στόχο την προετοιμασία των δεδομένων για Εξόρυξη Γνώσης και την εφαρμογή αλγορίθμων Μηχανικής Μάθησης
- Να κατανοήσουν τον τρόπο λειτουργίας διάφορων αλγορίθμων κατηγοριοποίησης (εποπτευμένη μάθηση), συσταδοποίησης (μη εποπτευμένη μάθηση) και εξαγωγής κανόνων συσχέτισης
- Να εκτιμούν την απόδοση των αλγορίθμων εξόρυξης γνώσης και μηχανικής μάθησης χρησιμοποιώντας τις κατάλληλες τεχνικές επικύρωσης και να αξιολογούν την παραγόμενα αποτελέσματα με στόχο τη λήψη αποφάσεων

5. ΠΕΡΙΓΡΑΜΜΑΤΑ ΜΑΘΗΜΑΤΩΝ

- Να σχεδιάζουν και να υλοποιούν Αποθήκες Δεδομένων και εφαρμόζουν αναλυτική επεξεργασία τύπου OLAP και αλγορίθμους εξόρυξης γνώσης σε αυτές

■ Γενικές Ικανότητες

- Αναζήτηση, ανάλυση και σύνθεση δεδομένων και πληροφοριών, με τη χρήση και των απαραίτητων τεχνολογιών
- Λήψη αποφάσεων
- Αυτόνομη ή και ομαδική εργασία
- Προαγωγή της ελεύθερης, δημιουργικής και επαγωγικής σκέψης

(γ) ΠΕΡΙΕΧΟΜΕΝΟ ΜΑΘΗΜΑΤΟΣ

- Εισαγωγή στην εξόρυξη γνώσης και τη μηχανικής μάθηση.
- Προετοιμασία δεδομένων (Καθαρισμός δεδομένων, διαχείριση απύσων τιμών, επιλογή χαρακτηριστικών, διακριτοποίηση, διαχείριση ανισοκατανομών σε προβλήματα κατηγοριοποίησης κ.α.)
- Εισαγωγή στην κατηγοριοποίηση/εποπτευμένη μάθηση, Κατηγορίες αλγορίθμων κατηγοριοποίησης, Αλγόριθμοι που βασίζονται σε πιθανότητες (π.χ. naïve bayes), Αλγόριθμοι που βασίζονται στον διαχωρισμό του χώρου (π.χ. Δένδρα αποφάσεων), Αλγόριθμοι που βασίζονται στην ομοιότητα/απόσταση (κ. εγγύτεροι γείτονες), αποτελεσματική αναζήτηση εγγύτερων γειτόνων μέσω δεικτοδότησης δεδομένων (π.χ. kd-tree), τεχνικές μείωσης δεδομένων εκπαίδευσης.
- Εισαγωγή στην συσταδοποίησης/μη εποπτευμένη μάθηση, Αλγόριθμοι Συσταδοποίησης: Ο αλγόριθμος κ μέσων (k-means) και οι παραλλαγές του, Ιεραρχική συσταδοποίηση, συσταδοποίηση βάσει πυκνότητας (DBSCAN), τεχνικές προσδιορισμού παραμέτρων
- Μέτρηση της απόδοσης αλγορίθμων κατηγοριοποίησης και συσταδοποίησης
- Κανόνες συσχέτισης, Ο αλγόριθμος Apriori για την ανακάλυψη κανόνων συσχέτισης, μέτρα αξιολόγησης των κανόνων συσχέτισης
- Επεξεργασία τύπου OLTP και OLAP, σχεδίαση και υλοποίηση Αποθήκης Δεδομένων, σχήμα αστέρα και χιονονιφάδας, διαδικασίες Extract-Transform-Load (ETL), πολυδιάστατοι κύβοι δεδομένων, αίτηματα OLAP και αλγόριθμοι εξόρυξης γνώσης σε κύβους δεδομένων.

(δ) ΔΙΔΑΚΤΙΚΕΣ και ΜΑΘΗΣΙΑΚΕΣ ΜΕΘΟΔΟΙ – ΑΞΙΟΛΟΓΗΣΗ

■ Τρόπος Παράδοσης

Θεωρητική από έδρας διδασκαλία με συζήτηση και ενεργή συμμετοχή των φοιτητών.

■ Χρήση Τεχνολογιών Πληροφορικής και Επικοινωνιών

Παρουσιάσεις μέσω projector.
Ηλεκτρονικές ασκήσεις αυτό-αξιολόγησης και υποστήριξη μαθησιακής διαδικασίας μέσω της ηλεκτρονικής πλατφόρμας eClass.

■ Οργάνωση Διδασκαλίας

Δραστηριότητα	Φόρτος Εργασίας Εξαμήνου
Διαλέξεις	39
Φροντιστηριακή διδασκαλία	13
Αυτοτελής μελέτη	98
Σύνολο μαθήματος	150

■ Αξιολόγηση Φοιτητών

Το μάθημα αξιολογείται με γραπτή τελική εξέταση δίωρης διάρκειας, και πιθανή διαδικασία διαρκούς αξιολόγησης κατά την κρίση του διδάσκοντα. Η ακριβής διαδικασία αξιολόγησης ανακοινώνεται στους φοιτητές και αναρτάται στο eclass στην αρχή του εξαμήνου.

(Ε) ΣΥΝΙΣΤΩΜΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

■ Προτεινόμενη Βιβλιογραφία

1. M. Dunham, *Data Mining*, Εκδόσεις νέων τεχνολογιών IKE
2. Tan Pang - Ning, Steinbach Michael, Kumar Vipin, Βερύκιος Βασίλειος (επιμέλεια), *Εισαγωγή στην εξόρυξη δεδομένων*, Εκδόσεις: Α. ΤΖΙΟΛΑ & ΥΙΟΙ Α.Ε., 2018
3. Mohammed J. Zaki, Wagner Meira JR, *Εξόρυξη και Ανάλυση Δεδομένων: Βασικές Έννοιες και Αλγόριθμοι*, Εκδόσεις Κλειδάριθμος
4. Anand Rajaraman, Jeffrey David Ullman, *Εξόρυξη από Μεγάλα Σύνολα Δεδομένων*, Εκδόσεις Νέων Τεχνολογιών IKE
5. Αλ. Νανόπουλος, Γ. Μανωλόπουλος, *Εισαγωγή στην Εξόρυξη Δεδομένων και τις Αποθήκες Δεδομένων*, Εκδόσεις Νέων Τεχνολογιών, 2008