

**ΤΙΤΛΟΣ ΜΑΘΗΜΑΤΟΣ:** Εξόρυξη Γνώσης από Δεδομένα-Μεθοδολογίες και Εφαρμογές

**ΕΚΠΑΙΔΕΥΤΗΣ:** Μανώλης Μαραγκουδάκης, Επίκουρος Καθηγητής,  
Πανεπιστημίου Αιγαίου

**ΣΤΟΧΟΣ ΜΑΘΗΜΑΤΟΣ:** Το μάθημα έχει στόχο να εισάγει τους καταρτιζόμενους σε βασικές μεθόδους της Εξόρυξης Γνώσης από Δεδομένα, μιας σχετικά νέας επιστημονικής περιοχής που προέκυψε από την ανάγκη επεξεργασίας μεγάλου όγκου δεδομένων και κατάφερε να εξελιχθεί σε ένα από τα σημαντικότερα ερευνητικά θέματα στην περιοχή των Πληροφοριακών Συστημάτων. Η ανάγκη για βελτίωση της ποιότητας των παρεχόμενων πληροφοριών καθώς επίσης και της εξαγωγής χρήσιμων συμπερασμάτων, επισπεύσανε την εξεύρεση τεχνικών διαχείρισης και επεξεργασίας δεδομένων. Η συνεισφορά των διαφόρων αυτών τεχνικών στην ανάλυση μεγάλων βάσεων δεδομένων σε σχέση με τις παραδοσιακές στατιστικές μεθόδους αναγνωρίζεται όλο και περισσότερο. Επομένως, ενώ οι στατιστικές δοκιμές απαιτούν πραγματοποίηση ελέγχων στατιστικών υποθέσεων αναφορικά με έναν πληθυσμό, οι τεχνικές εξόρυξης δεδομένων παρέχουν με αυτοματοποιημένο τρόπο επιβεβαιωμένες σχέσεις με τη μορφή κανόνων για μεγάλες βάσεις δεδομένων.

. Ιδιαίτερη βαρύτητα δίνεται στις αρχές και στις μεθόδους κατηγοριοποίησης, συσταδοποίησης και ανακάλυψης κανόνων συσχέτισης. Στο υλικό του μαθήματος δίνονται συστηματικά οι βασικές αρχές που θα πρέπει να ακολουθούνται σε μια διαδικασία εξόρυξης γνώσης από δεδομένα και οι λανθασμένες ενέργειες που θα πρέπει να αποφεύγονται στα διάφορα στάδια αυτής, ώστε να γίνει ευκολότερη κάθε προσπάθεια ενασχόλησης με τα σχετικά θέματα ή τουλάχιστον ελέγχου της ορθότητας ορισμένων βασικών επιλογών που γίνονται στα πλαίσια αυτής.

**ΧΡΗΣΙΜΟΤΗΤΑ ΤΟΥ ΜΑΘΗΜΑΤΟΣ:** Κατανόηση των μεθόδων αυτών θα δώσει στον καταρτιζόμενο ένα σημαντικό εργαλείο ιδιαίτερα χρήσιμο στη διερεύνηση και ερμηνεία των συσχετίσεων που ενυπάρχουν σε μια συλλογή δεδομένων, την δημιουργία ευφών διαδικασιών πρόβλεψης μιας μελλοντικής τάσης ή την εύρεση κρυμμένων αλλά ουσιαστικών συσχετίσεων. Η εξόρυξη γνώσης από δεδομένα μπορεί να καταστεί ένα ιδιαίτερα χρήσιμο εργαλείο στις σύγχρονες επιχειρήσεις και οργανισμούς, καθώς η ολοένα αυξανόμενη ροή δεδομένων συνιστά

επιτακτική την ανάγκη για άμεση και αποτελεσματική εξαγωγή χρήσιμης και δομημένης πληροφορίας από αυτά.

**ΣΕ ΠΟΙΟΥΣ ΑΠΕΥΘΥΝΕΤΑΙ ΤΟ ΜΑΘΗΜΑ:** Το μάθημα απευθύνεται στους ερευνητές που επιθυμούν να ασχοληθούν με τα Πληροφοριακά Συστήματα ή εμπλέκονται σ' αυτή στα πλαίσια της επαγγελματικής τους ή ακαδημαϊκής ενασχόλησης. Στελέχη της διοίκησης του δημόσιου ή του ιδιωτικού τομέα θα βρουν λύσεις στην ερευνητική και επαγγελματική τους δραστηριότητα που απαιτείται για την ορθολογική λήψη αποφάσεων.

Επισημαίνεται ότι στο υλικό του μαθήματος έχει μεταφερθεί η εμπειρία του συγγραφέα από την μακρόχρονη ενασχόλησή του με την εξόρυξη γνώσης από ερευνητική εργασία του στον Πανεπιστημιακό χώρο.

## **ΠΕΡΙΕΧΟΜΕΝΑ ΤΟΥ ΜΑΘΗΜΑΤΟΣ**

### ***ΕΝΟΤΗΤΑ 1: ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΣΤΗΝ ΕΞΟΡΥΞΗ ΓΝΩΣΗΣ***

#### **ΚΕΦΑΛΑΙΟ 1: Εισαγωγή στην Εξόρυξη Γνώσης**

Στο κεφάλαιο αυτό δίνονται εισαγωγικές έννοιες για τον ορισμό, το περιεχόμενο, το πεδίο εφαρμογής και τα στάδια της εξόρυξης γνώσης, προκειμένου αυτά να αποτελέσουν τη βάση καλύτερης κατανόησης των επόμενων κεφαλαίων.

#### **ΚΕΦΑΛΑΙΟ 2: Στρατηγικές Εξόρυξης Γνώσης**

Στο συγκεκριμένο κεφάλαιο δίνονται οι βασικές κατηγορίες που περιέχει ο τομέας της Εξόρυξης Γνώσης, αναφέρονται τα μέτρα αξιολόγησης των επιδόσεων συστημάτων εξόρυξης και τέλος, μια βασική αναφορά στις κοινωνικές επιπτώσεις.

#### **ΚΕΦΑΛΑΙΟ 3: Τεχνικές Οπτικοποίησης**

Στο κεφάλαιο αυτό αναφέρονται οι διαδικασίες για την πρώτη ανάλυση των δεδομένων με χρήση γραφικών εργαλείων, μια διαδικασία που εξοικειώνει το χρήστη με τις ιδιότητες των δεδομένων και παράγει μια αρχική αξιολόγηση των παραμέτρων που θα καθορίσουν τη διαδικασία εξόρυξης.

## ***ΕΝΟΤΗΤΑ 2: ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗ***

### ΚΕΦΑΛΑΙΟ 4: Βασική Κατηγοριοποίηση

Στο κεφάλαιο αυτό περιγράφονται αναλυτικά οι αλγόριθμοι και οι μεθοδολογίες που επιτρέπουν σε ένα σύστημα εξόρυξης γνώσης από δεδομένα να πραγματοποιήσει κατηγοριοποίηση ή πρόβλεψη για μια μελλοντική, άγνωστη κατάσταση με βάση την εμπειρία από προγενέστερα δεδομένα. Δίνονται παραδείγματα από εφαρμογές του πραγματικού κόσμου.

## ***ΕΝΟΤΗΤΑ 2: ΣΥΣΤΑΔΟΠΟΙΗΣΗ***

### ΚΕΦΑΛΑΙΟ 5 Αλγόριθμοι Συσταδοποίησης

Στο κεφάλαιο αυτό αναπτύσσεται με λεπτομέρεια η μεθοδολογία των επικρατέστερων αλγορίθμων συσταδοποίησης δεδομένων, μιας διαδικασίας που ομαδοποιεί αυτόματα δεδομένα με βάση τα χαρακτηριστικά τους και διευκολύνει τους ερευνητές στο να αναγνωρίζουν όμοιες συμπεριφορές και τάσεις σε μια πληθώρα δεδομένων και πληροφορίας. Όπως και στο προηγούμενο κεφάλαιο, αναφέρονται με λεπτομέρεια παραδείγματα από εφαρμογές του πραγματικού κόσμου.

## ***ΕΝΟΤΗΤΑ 3: ΑΝΑΚΑΛΥΨΗ ΚΑΝΟΝΩΝ ΣΥΣΧΕΤΙΣΗΣ***

### ΚΕΦΑΛΑΙΟ 6 Αλγόριθμοι Ανακάλυψης Κανόνων Συσχέτισης

Στο κεφάλαιο αυτό δίνονται αναλυτικά οι επικρατέστεροι αλγόριθμοι ανακάλυψης κανόνων συσχέτισης από δυναμικά δεδομένα, μιας διαδικασίας από το χώρο της εξόρυξης, η οποία αποτυπώνει με γραφικό τρόπο την συσχέτιση μεταξύ παραγόντων και χαρακτηριστικών των δεδομένων και συμβάλλει στην δημιουργία συστημάτων σύστασης (recommendation systems), μιας νέας τάσης υπηρεσιών με μεγάλη απήχηση στο ευρύ κοινό (π.χ. η υπηρεσία σύστασης προϊόντων Amazon). Αναφέρουμε επίσης ότι γίνεται αναφορά και μελέτη πραγματικών εφαρμογών από την καθημερινότητα.