



Πανεπιστήμιο Πειραιώς
Τμήμα Πληροφορικής

Ακαδημαϊκό έτος 2009-10 (χειμερινό εξάμηνο)

ΑΠΟΘΗΚΕΣ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΕΞΟΡΥΞΗ ΓΝΩΣΗΣ (Data Warehousing and Data Mining – DWDM)

διδάσκων: Αν. Καθ. Γιάννης Θεοδωρίδης (ytheod@unipi.gr)
υπεύθυνος εργαστηρίου: Δρ. Νίκος Πελέκης (npelekis@unipi.gr)
βοηθοί εργαστηρίου: Νίκος Γιατράκος (ngiatrak@unipi.gr),
Δέσποινα Κοπανάκη (dkorpanak@unipi.gr), Γεράσιμος Μαρκέτος (marketos@unipi.gr)

ΣΥΝΟΠΤΙΚΗ ΠΕΡΙΓΡΑΦΗ ΜΑΘΗΜΑΤΟΣ

Αντικείμενο του μαθήματος

Σκοποί του μαθήματος είναι (α) η μελέτη τεχνικών εξόρυξης γνώσης από δεδομένα και (β) η παρουσίαση εφαρμογών και σεναρίων χρήσης αποτελεσμάτων εξόρυξης γνώσης, με έμφαση σε εξόρυξη από μεγάλες βάσεις δεδομένων.

Τρόπος διεξαγωγής του μαθήματος

Το μάθημα αποτελείται από διαλέξεις¹ σε θέματα όπως (ενδεικτικά):

- Αποθήκες δεδομένων (κύβοι, πολυδιάστατο μοντέλο δεδομένων) και λειτουργίες OLAP,
- Βασικές τεχνικές data mining (κατηγοριοποίηση, συσταδοποίηση, εξόρυξη κανόνων συσχέτισης),
- Προηγμένες τεχνικές data mining (εξόρυξη χρονικής / χρονικής γνώσης, εξόρυξη στο web),

Παράλληλα, θα υπάρχουν εργαστηριακές ώρες² για βασική εκμάθηση των δυνατοτήτων σε Business Intelligence (OLAP / data warehousing / data mining) των εμπορικών ΣΔΒΔ SQL Server και Oracle, καθώς και του εργαλείου data mining ανοικτού λογισμικού WEKA.

¹ Οι διαλέξεις θα γίνονται κάθε Πέμπτη 3-6μμ στο Π. Αμφ. – έναρξη στις 15.10.09

² Τα εργαστήρια θα γίνονται κάθε Παρασκευή 2-4μμ στην αίθ. 206 – έναρξη στις 30.10.09

Τρόπος βαθμολόγησης

Κατά τη διάρκεια του μαθήματος θα δοθούν δύο εργασίες (μία υποχρεωτική και μία προαιρετική) με συμμετοχή 50-50 στον τελικό βαθμό:

- 1^η εργασία (υποχρεωτική – σε ομάδες των 2-3 ατόμων): Επιλογή μιας μεγάλης βάσης δεδομένων, προετοιμασία δεδομένων για OLAP και data mining, επίδειξη αποτελεσμάτων OLAP και data mining, ερμηνεία αποτελεσμάτων, συγγραφή σχετικής αναφοράς.
- 2^η εργασία (προαιρετική - ατομική): Συμμετοχή στην ανάπτυξη ενός DW/DM project, είτε πρακτικά (προγραμματιστικά) είτε θεωρητικά (π.χ. επισκόπηση σχετικής βιβλιογραφίας ή αποτίμηση ενός αλγορίθμου).

Στο τέλος του εξαμήνου, υπάρχει γραπτή εξέταση (με ανοικτά βιβλία-σημειώσεις) μόνο για όσους δεν επέλεξαν να υλοποιήσουν τη 2^η εργασία (δηλαδή, συμμετοχή 50% στον τελικό βαθμό).

Προαπαιτούμενες γνώσεις

Το μάθημα προϋποθέτει γνώσεις Βάσεων Δεδομένων και Στατιστικής, όπως μοντελοποίηση σχεσιακών βάσεων δεδομένων, γλώσσα SQL, θέματα ακεραιότητας δεδομένων, βασικά στοιχεία στατιστικής ανάλυσης δεδομένων.

Προτεινόμενη βιβλιογραφία

- Διδακτικά βιβλία:
 - Dunham M. “Data Mining – Εισαγωγικά και Προηγμένα Θέματα Εξόρυξης Γνώσης από Δεδομένα”. Εκδόσεις Νέων Τεχνολογιών, 2004.
 - Μανωλόπουλος Ι., Νανόπουλος Αλ. “Εισαγωγή στην Εξόρυξη Δεδομένων και τις Αποθήκες Δεδομένων”. Εκδόσεις Νέων Τεχνολογιών, 2009.
 - Roiger R.G., Geatz M.W. “Εξόρυξη Πληροφορίας – Ενας Εισαγωγικός Οδηγός με Παραδείγματα”. Εκδόσεις Κλειδάριθμος, 2008.
- Διαφάνειες διδάσκοντα (διαθέσιμες στην ιστοσελίδα του μαθήματος)

Περισσότερες πληροφορίες

- Ιστοσελίδα μαθήματος: <http://infolab.cs.unipi.gr/courses/dwdm.html>

Καλό εξάμηνο !