

ΕΠΛ660: ΑΝΑΚΤΗΣΗ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΜΗΧΑΝΕΣ ΑΝΑΖΗΤΗΣΗΣ

Διδακτικές
Μονάδες 8

Διδάσκων Γιώργος Πάλλης

Στόχοι και Μεθοδολογία

Το μάθημα αυτό παρέχεται στο πλαίσιο του μεταπτυχιακού προγράμματος του Τμήματος Πληροφορικής στο Πανεπιστήμιο Κύπρου. Το μάθημα αποτελεί μια εισαγωγή στην θεωρία και τις τεχνικές της Ανάκτησης Πληροφοριών και των Μηχανών Αναζήτησης.

Συγκεκριμένα, θα εξετασθούν:

- Μπούλειος Ανάκτηση Πληροφοριών
- Κωδικοποίηση κειμένου, λημματοποίηση, στελέχωση κειμένων
- Λεξικά και ανάκτηση ανεκτική σε σφάλματα
- Κατασκευή και συμπίεση ευρετηρίων
- Διαβάθμιση όρων
- Ανάκτηση διανυσματικού χώρου
- Αξιολόγηση ανάκτησης πληροφοριών
- Μηχανισμοί ανάδρασης και διαστολή επερωτήσεων
- Ταξινόμηση κειμένου και απλοϊκές τεχνικές Bayes
- Ταξινόμηση διανυσματικού χώρου
- Επίπεδη ομαδοποίηση
- Ιεραρχική ομαδοποίηση
- Βασικές έννοιες αναζήτησης στον Ιστό
- Ανάκτηση πληροφορίας σε ημι-δομημένα δεδομένα
- Αναζήτηση λογισμικού σε Υπολογιστικές Νεφέλες
- Ιχνηλασία και ευρετηριασμός Ιστού
- Ανάλυση υπερσυνδέσμων

Το μάθημα αποτελείται από 3 ώρες διαλέξεων και μία ώρα εβδομαδιαίου εργαστηρίου-φροντιστηρίου. Η ύλη των διαλέξεων θα συνοδεύεται από ασκήσεις και εργασίες.

Προαπαιτούμενα

Το μάθημα είναι ανοικτό σε όλους τους μεταπτυχιακούς φοιτητές Πληροφορικής και Μηχανικής Υπολογιστών, όπως και σε τελειόφοιτους φοιτητές Πληροφορικής. Για την απρόσκοπτη παρακολούθηση του μαθήματος ότι οι προπτυχιακοί φοιτητές πρέπει να έχουν παρακολουθήσει επιτυχώς τα μαθήματα Δομών Δεδομένων και Αλγορίθμων, Αντικειμενοστρεφούς Προγραμματισμού, Πιθανοτήτων και Στατιστικής, Γραμμικής Άλγεβρας και Βάσεων Δεδομένων.

Αξιολόγηση και Βαθμολογία

Η αξιολόγηση των φοιτητών θα βασισθεί σε προγραμματιστικές εργασίες, εργασία εξαμήνου και στην τελική εξέταση. Ο τελικός βαθμός μαθήματος διαμορφώνεται με βάση τα ακόλουθα ποσοστά:

- Εργασία Εξαμήνου 25%
- Εργασίες (προγραμματιστικές) 15% (κατά την υποβολή στο εργαστήριο)
- Γραπτή Ενδιάμεση Εξέταση 20%
- Γραπτή Τελική Εξέταση 40%

Σημειώνεται ότι τυχόν αποτυχία του φοιτητή να συγκεντρώσει πάνω από 45/100 στο σταθμισμένο μέσο όρο ενδιάμεσης και τελικής εξέτασης, ισοδυναμεί αυτομάτως με αποτυχία στο μάθημα, ανεξαρτήτως της επιδόσεως στις εργασίες. Σε τέτοια περίπτωση, ο τελικός βαθμός του μαθήματος που θα δοθεί στον φοιτητή θα είναι ίσος με τον σταθμισμένο μέσο όρο ενδιάμεσης και της τελικής εξέτασης.

Βιβλιογραφία

Θα χρησιμοποιηθεί το ακόλουθο σύγγραμμα:

Introduction to Information Retrieval, by Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze, Cambridge University Press, 2008. Το βιβλίο αυτό υπάρχει σε ηλεκτρονική μορφή στο Διαδίκτυο, μαζί με χρήσιμο μαθησιακό υλικό.

Χρήσιμο μπορεί να σας φανεί και το βιβλίο Modern Information Retrieval των Baeza-Yates και Ribeiro-Neto.

Θα χρησιμοποιηθεί επίσης υλικό και συλλογή άρθρων που θα δοθούν στη διάρκεια του μαθήματος.