



بازیابی پیشرفته اطلاعات

Modern Information Retrieval

طراح درس: حسن ابوالحسنی

پیش نیاز: -

واحد: ۳

گرایش: نرم افزار

مقطع: کارشناسی

کلیات

هدف این درس، آشنا کردن دانشجویان کارشناسی با مفاهیم جدید در حوزه بازیابی اطلاعات با تاکید بر موتورهای جستجوی وب است. یک بخش عمده این درس معطوف به سه مدل بازیابی پایه یعنی مدل بولی، مدل برداری و مدل احتمالی در بازیابی اطلاعات می‌باشد. در کنار آشنایی با این مدلها داده‌ساختارهای اساسی برای پیاده‌سازی یک موتور جستجو نیز مورد بحث قرار می‌گیرد. ارزیابی سیستم‌های بازیابی اطلاعات، تغییر و گسترش پرسمان، بازیابی ایکس ام ال، دسته بندی اسناد و خوشه بندی اسناد از مطالب دیگر درس است. یک بخش مهم دیگر این درس مطالعه خصوصیات وب، خزنده‌های وب و مفهوم تحلیل لینک است.

ریز مواد

۱. مقدمه - معرفی بازیابی اطلاعات و مدل بولی (۱ جلسه)
۲. مراحل ساخت دیکشنری لغات (۱ تا ۲ جلسه)
۳. بازیابی منعطف (۱ تا ۲ جلسه)
۴. ساخت اندیس (۱ تا ۲ جلسه)
۵. فشرده سازی اندیس (۱ تا ۲ جلسه)
۶. مدل برداری (۲ تا ۳ جلسه)
۷. نحوه ارزیابی سیستمهای بازیابی اطلاعات (۱ تا ۲ جلسه)
۸. توسعه پرسمان (۱ تا ۲ جلسه)
۹. بازیابی اسناد ایکس ام ال (۱ تا ۲ جلسه)
۱۰. مدل احتمالی (۲ تا ۳ جلسه)
۱۱. دسته بندی اسناد (۲ تا ۳ جلسه)
۱۲. خوشه بندی اسناد (۲ تا ۳ جلسه)
۱۳. مشخصه‌های وب (۱ تا ۲ جلسه)
۱۴. معماری خزنده‌های وب (۲ تا ۳ جلسه)
۱۵. تحلیل پیوند در وب (۲ تا ۳ جلسه)

آزمون - تمرین - تکلیف

- آزمون: آزمونهای میان ترم و پایان ترم (۷۰٪ کل نمره)
- تمرین: سه تا چهار تمرین عملی در طول نیمسال تحویل داده می شوند. (۲۰٪ کل نمره)
- تکلیف: چهار تا پنج تکلیف در طول نیمسال تحویل داده می شوند. (۱۰٪ کل نمره)

مراجع اصلي

1. Introduction to information retrieval, Cambridge University Press, 2008.
2. Modern Information Retrieval, Baeza Yates, 1999.