



**LUNDS UNIVERSITET**  
Lunds Tekniska Högskola

*Kursplan för*

## **Webb-intelligens och informationssökning Web Intelligence and Information Retrieval**

**EITN01, 7,5 högskolepoäng, A (Avancerad nivå)**

**Gäller för:** Läsåret 2012/13

**Beslutad av:** Utbildningsnämnd 1

**Beslutsdatum:** 2012-03-19

### **Allmänna uppgifter**

**Valfri för:** C4, D4, D4-ks

**Undervisningsspråk:** Kursen kan komma att ges på engelska

### **Syfte**

Kursens syfte är att öka förståelsen för metoder för informationssökning (Information Retrieval), strukturering, och kunskapsextraktion (text mining) speciellt från Internet-baserade källor.

### **Mål**

*Kunskap och förståelse*

För godkänd kurs skall studenten

på egen hand kunna:

- identifiera och formulera problem inom området Webb-Intelligens
- klassificera problemens svårighetsgrad i förhållande till den egna kunskapsnivån
- förstå och använda metoder från området informationssökning
- analysera informationskällor för att skapa strukturerad information

*Färdighet och förmåga*

För godkänd kurs skall studenten

- kunna implementera intelligenta applikationer för kunskapsextraktion av medelhög komplexitet, med utnyttjande av metodologi från informationssökning

*Värderingsförmåga och förhållningssätt*

För godkänd kurs skall studenten

- visa prov på förmåga att hantera för studenten nya metoder och resultat

## Kursinnehåll

*Informationssökning*: Basala begrepp för sökning och ranking, vektormodeller, tf-idf baserad relevans. System för informationssökning.

*Sökspråk*: Olika sökspråk för sökning i strukturerade databaser presenteras.

*Strukturerad information*: Indexering, sökning och relevansranking av sökresultat. Exemplifieras med hjälp av sökningar i strukturerade databaser (SRU/CQL).

*"Feature extraction"*: Extrahera egenskaper för dokument.

Grundläggande metoder för kunskapsextraktion och klassificering (som neurala nätverk, Support Vector Machines m fl) gås igenom och experimenteras med. Utnyttja extraherade egenskaper för att implementera ämnes-klassifikation för textdokument.

*Prestanda*: Prestandamätetal som precision och "recall".

## Kursens examination

**Betygsskala**: TH

**Prestationsbedömning**: Tentamen samt godkända laborationer och inlämningsuppgifter.

### Delmoment

**Kod**: 0113. **Benämning**: Laborationer.

Antal högskolepoäng: 4. **Betygsskala**: UG. **Prestationsbedömning**: Godkända laborationer.

**Kod**: 0213. **Benämning**: Inlämningsuppgifter.

Antal högskolepoäng: 0. **Betygsskala**: UG. **Prestationsbedömning**: Godkända inlämningsuppgifter.

**Kod**: 0313. **Benämning**: Tentamen.

Antal högskolepoäng: 3,5. **Betygsskala**: TH. **Prestationsbedömning**: Godkänd tentamen.

## Antagningsuppgifter

**Förkunskapskrav**:

- EDA011 eller EDA016 Programmeringsteknik

**Förutsatta förkunskaper**: FMA420 Linjär Algebra.

**Begränsat antal platser**: Nej

**Kursen kan ställas in**: Om färre än 8 anmälda.

**Kursen överlappar följande kurser**: EIT031

## Kurslitteratur

- Baeza-Yates, R, Ribeiro-Neto, B: Modern Information Retrieval.
- Addison-Wesley 1999. ISBN: 0-201-39829-X.
- Artiklar och document från Webben.
- Kursnoter och laborationer.

## Kontaktinfo och övrigt

**Kursansvarig**: Universitetslektor Anders Ardo, Anders.Ardo@eit.lth.se

**Hemsida**: <http://www.eit.lth.se/kurs/eitn01>

**Övrig information:** Vid färre än 16 deltagare kan kursen komma att ges med reducerad undervisning och större inslag av självstudier.