



DD2475 Informationssökning

9,0 hp

Information Retrieval

När kurs inte längre ges har student möjlighet att examineras under ytterligare två läsår.

Fastställande

Kursplan för DD2475 gäller från och med HT10

Betygsskala

A, B, C, D, E, FX, F

Utbildningsnivå

Avancerad nivå

Huvudområden

Datalogi och datateknik

Undervisningsspråk

Undervisningsspråk anges i kurstillfällesinformationen i kurs- och programkatalogen.

Lärandemål

Efter genomförd kurs ska du kunna:

- förklara koncepten indexering, vokabulär, normalisering och kodbok i informationssökning

- definiera en boolesk modell och en vektorrumsmodell, och förklara skillnaden mellan dem
- förklara skillnaderna mellan klassificering och klustring
- diskutera skillnaderna mellan olika klassificerings- och klustringsmetoder
- välja en lämplig klassificerings- eller klustringsmetod beroende på det aktuella problemets begränsningar
- implementera klassificering i en boolesk modell och i en vektorrumsmodell
- implementera en grundläggande klustringsmetod
- redogöra för en grundläggande spektralmetod
- utvärdera informationssökningsalgoritmer, och redogöra för svårigheter med utvärdering
- förklara grunderna i XML och internetsökning.

Kursinnehåll

Grundläggande och avancerade tekniker för informationssökning; informationsextraktion; effektiv textindexering; indexering av icke-textdata; booleska modeller och vektorrumsmodeller för sökning; utvärderings- och gränssnittsfrågor; XML, internetsökmotorers uppbyggnad; klustring, klassificering; spektralmetoder, random indexing; sökning i stora mängder data.

Särskild behörighet

För fristående kursstuderande krävs 90 högskolepoäng varav 45 högskolepoäng inom matematik eller informationsteknik. Dessutom krävs engelska B eller motsvarande.

Kurslitteratur

C. D. Manning, P. Raghavan and H. Schütze: Introduction to Information Retrieval, Cambridge University Press, 2008.

Examination

- LAB1 - Laborationer, 3,0 hp, betygsskala: P, F
- LAB2 - Projekt, 3,0 hp, betygsskala: A, B, C, D, E, FX, F
- TEN1 - Tentamen, 3,0 hp, betygsskala: A, B, C, D, E, FX, F

Examinator beslutar, baserat på rekommendation från KTH:s samordnare för funktionsnedsättning, om eventuell anpassad examination för studenter med dokumenterad, varaktig funktionsnedsättning.

Examinator får medge annan examinationsform vid omexamination av enstaka studenter.

I denna kurs tillämpas skolans hederskodex, se: <http://www.kth.se/csc/student/hederskodex>.

Övriga krav för slutbetyg

Deltagarna i kursen förväntas medverka aktivt i kursens olika moment, i synnerhet vid övningar och laborationer. Dessutom betonas:

- träning i att självständigt inhämta nödvändig kunskap
- träning i muntlig och skriftlig redovisning.

Examinationen sker genom tentamen (TEN1; 3,0 hp), datorlaborationer (LAB1; 3,0 hp) och en projektuppgift som redovisas skriftligt och muntligt (LAB2; 3,0 hp).

Etiskt förhållningssätt

- Vid grupparbete har alla i gruppen ansvar för gruppens arbete.
- Vid examination ska varje student ärligt redovisa hjälp som erhållits och källor som använts.
- Vid muntlig examination ska varje student kunna redogöra för hela uppgiften och hela lösningen.