HÖGSKOLAN I BORÅS

Klassifikation och informationsutvinning
Classification and Data Mining

7,5 högskolepoäng
7,5 credits

Ladokkod: 32IKI1
Version: 5.0
Fastställd av: Utbildningsutskottet 2012-04-18
Gäller från: VT 2012

Nivå: Avancerad nivå
Huvudområde (successiv fördjupning): Biblioteks- och informationsvetenskap (A1N)
Utbildningsområde: Övrigt
Ämnesgrupp: Biblioteks- och informationsvetenskap
Förkunskapskrav: Kandidatexamen 180 högskolepoäng i biblioteks- och informationsvetenskap.
Betygsskala: Underkänd, Godkänd eller Väl godkänd

Innehåll

• Allmän klassifikationsteori
• Övervakad maskininlärning
• Oövervakad maskininlärning
• Val av särdrag
• Utvärdering av klassifikation
• Informationsvisualisering för IR

Mål

Kursen syftar till att illustrera utvecklingsriktningar inom IR-forskningen som är av vikt för digitala bibliotek.
Efter avslutad kurs ska studenterna kunna:

- Analysera sambandet mellan informationsutvinning och IR i termer av innehållsrepresentation och innehållskategorisering.
- Resonera kring val av särdrag för informationsutvinning.
- Utvärdera informationsutvinning med hjälp av vanliga utvärderingsmått.
- Visa en djupare förståelse och praktisk kompetens inom informationsutvinning och informationsvisualisering.
- Förklara de grundläggande principerna bakom användningen av IR i ett digitalt bibliotek.

Undervisningsformer

Undervisningen bedrivs i form av laborationer, föreläsningar, praktiska genomgångar och övningar.

Undervisningen bedrivs normalt på svenska men undervisning på engelska kan förekomma.

Examinationsformer

Kursen examineras genom tentamina, rapporter och projektarbeten.

Studenten har rätt till minst fem (5) examinationstillfällen, varav minst tre (3) inom loppet av det första året.

Studentens rättigheter och skyldigheter vid examination är enligt riktlinjer och regelverk vid Högskolan i Borås.

Kurslitteratur och övriga läromedel

Baeza-Yates, R. & Ribeiro-Neto, B. (2011). Modern Information Retrieval: The Concepts and Technology Behind Search. Addison-Wesley: Harlow. (Kap. 2, s. 21-56; Kap. 8, s. 281-336.)

Manning, C.D., Raghavan, P. & Schütze, H. (2008). Introduction to information retrieval. Cambridge: Cambridge University Press. Kap. 8. (26 s.) [Tillgänglig elektroniskt]

Sebastiani, S. (2005). Text categorization. In Alessandro Zanasi (ed.), Text Mining and its Applications, WIT Press: Southampton, s. 109-129.

Stavrianou, A., Andritsos, P. & Nicoloyannis, N. (2007). Overview and Semantic Issues of Text Mining. SIGMOD Record, 36(3), s. 23-33.

Witten, I. & Frank, E. (2005). Data Mining: Practical Machine Learning Tools and Techniques 2nd. Ed.Morgan Kaufman Publishers: Amsterdam. s 3-17.

Studentinflytande och utvärdering

Studenternas synpunkter på kursen ska systematiskt och regelbundet inhämtas i muntlig och/eller skriftlig form. Studenterna ska delges resultatet av utvärderingen. Resultaten av utvärderingarna ska ligga till grund för kursens vidare utveckling.

Se vidare Policy för kursutvärdering, Högskolan i Borås 2005-06-07, dnr 56-02-10.

Övrigt

Kursen ingår i Masterprogram i Biblioteks- och informationsvetenskap, distansutbildning.