



INSTITUTIONEN FÖR FILOSOFI, LINGVISTIK OCH VETENSKAPSTEORI

LT2114 Praktisk grundläggande språkteknologi, 7,5 högskolepoäng

Practical natural language processing, 7.5 credits

Avancerad nivå / Second Cycle

Fastställande

Kursplanen är fastställd av Institutionen för filosofi, lingvistik och vetenskapsteori 2017-06-01 och senast reviderad 2019-03-14. Den reviderade kursplanen gäller från och med 2019-03-18, vårterminen 2019.

Utbildningsområde: Naturvetenskapligt 100 %

Ansvarig institution: Institutionen för filosofi, lingvistik och vetenskapsteori

Inplacering

Kan även ges som fristående kurs.

Kursen kan ingå i följande program: 1) Computer Science, Master's Programme (N2COS) och 2) Applied Data Science masterprogram (N2ADS)

Huvudområde

Språkteknologi

Fördjupning

A1N, Avancerad nivå, har endast kurs/er på grundnivå som förkunskapskrav

Förkunskapskrav

För tillträde till kursen krävs minst 7.5 poäng i programmering, t.ex. DIT948 Programmering, DIT012 Imperativ programmering med grundläggande objektorientering, DIT142 Funktionell programmering, LT2111 Introduktion till programmering, eller motsvarande.

Lärandemål

Efter godkänd kurs ska studenten kunna:

Kunskap och förståelse

- redogöra för grundläggande begrepp inom språkteknologin,
- i detalj förklara några vanliga algoritmer för språkteknologiska uppgifter, såsom textkategorisering eller ordklasstagning,
- redogöra för olika utvärderingsmetoder inom språkteknologin, samt när de är tillämpbara.

Färdigheter och förmåga

- använda ett språkteknologiskt programbibliotek för att utveckla praktiska språkteknologiska datorapplikationer,
- implementera metoder för grundläggande textbehandling,
- implementera utvärderingsprocedurer för språkteknologiska uppgifter.

Värderingsförmåga och förhållningssätt

- resonera om vilken språkteknologisk metod som är tillämpbar för att lösa ett givet språkteknologiskt problem,
- välja lämpliga mått för evaluering av en föreslagen lösning av en NLP-uppgift, samt motivera detta val.

Innehåll

Kursen ger en praktisk introduktion till olika problem inom språkteknologiområdet, samt några vanliga lösningar. Studenterna kommer att få praktisk programmeringsfärdighet genom att undersöka dessa tillämpningar. Kursen är uppdelad i fyra huvuddelar, en som täcker grundläggande begrepp och tre som tänker delområden inom språkteknologin: grundläggande textbehandling, syntaktisk och morfologisk analys, samt semantik och pragmatik. Det definitiva urvalet av ämnen som täcks i de fyra delområdena bestäms av den kursansvarige innan kursen startar. Det följande är ett exempel på ämnen som kan genomgå under kursen:

1. Grundläggande begrepp:

- grundläggande begrepp inom språkteknologin,
- sannolighetsteori och maskininlärning för språkteknologiska problem,
- vanliga utvärderingsmått, t.ex. korrekthet, precision och täckning.

2. Grundläggande textbehandling:

- korpusar och korpusannotering,
- grundläggande statistisk korpusanalys.

3. Syntax och morfologi:

- morfologisk analys (ordklassanalys) med hjälp av maskininlärning och finite-state-teknologi,
- syntaktisk analys med hjälp av regelbaserade och datadrivna metoder.

4. Semantik och pragmatik:

- ordbetydelsedisambiguering med hjälp av maskininlärning,
- textklassificering eller topic modeling med hjälp av maskininlärning.

Delkurser

1. Skriftlig tentamen (Written examination), 4 hp

Betygsskala: Väl godkänd (VG), Godkänd (G) och Underkänd (U)

2. Uppgifter (Assignments), 3,5 hp

Betygsskala: Väl godkänd (VG), Godkänd (G) och Underkänd (U)

Former för undervisning

Föreläsningar, övningar, datorlaborationer. Kursen har obligatoriska laborationstillfällen med krav på närvaro.

Undervisningsspråk: engelska

Former för bedömning

Kursen examineras med en kombination av skriftlig tentamen, inlämningsuppgifter, laborationer och aktivt deltagande vid övningar och laborationstillfällen.

Om student som underkänts två gånger på samma examinerande moment önskar byte av examinator inför nästa examinationstillfälle, ska sådan begäran inlämnas skriftligt till kursansvarig institution och bifallas om det inte finns särskilda skäl däremot (HF 6 kap § 22).

I det fall en kurs har upphört eller genomgått större förändringar ska studenten i normalfallet garanteras tillgång till minst tre provtillfällen (inklusive ordinarie provtillfälle) under en tid av åtminstone ett år med utgångspunkt i kursens tidigare uppläggning.

Betyg

På kursen ges något av betygen Väl godkänd (VG), Godkänd (G) och Underkänd (U).
För betyget Godkänd krävs:

- Godkänd skriftlig tentamen
- avslutade fullständiga inlämningsuppgifter
- aktivt deltagande vid laborationstillfällen

För betyget Väl godkänd krävs:

- Väl godkänd skriftlig tentamen
- avslutade fullständiga inlämningsuppgifter av god kvalitet
- aktivt deltagande vid laborationstillfällen

Kursvärdering

Studierektor eller motsvarande är ansvarig för att resultat av studenternas kursvärderingar sammanställs, tillgängliggörs och ligger till underlag för kursens utveckling.

Resultatet och eventuella förändringar i kursens upplägg ska förmedlas både till de studenter som genomförde värderingen och till de studenter som ska påbörja kursen.