



# INSTITUTIONEN FÖR FILOSOFI, LINGVISTIK OCH VETENSKAPSTEORI

## LT2822 Maskininläring för statistisk datalingvistik: inledning, 7,5 högskolepoäng

Machine learning for statistical NLP: introduction, 7.5 credits

*Avancerad nivå / Second Cycle*

---

### Fastställande

Kursplanen är fastställd av Institutionen för filosofi, lingvistik och vetenskapsteori 2023-11-13 att gälla från och med 2024-01-15, vårterminen 2024.

*Utbildningsområde:* Naturvetenskapligt 100 %

*Ansvarig institution:* Institutionen för filosofi, lingvistik och vetenskapsteori

### Inplacering

Kan också ges som fristående kurs.

#### *Huvudområde*

Språkteknologi

#### *Fördjupning*

A1F, Avancerad nivå, har kurs/er på avancerad nivå som förkunskapskrav

### Förkunskapskrav

För tillträde till kursen krävs antingen godkänt resultat på:

- programmering, 7,5hp och
- formell lingvistik, 7,5hp.

### Lärandemål

Efter godkänd kurs ska studenten kunna:

#### *Kunskap och förståelse*

- redogöra för användningen av språkliga data i tillämpningar inom språkteknologi,
- visa förståelse för grundläggande metodologi inom, och likheter mellan, språkteknologi och maskininläring,

- visa förståelse för den matematiska grunden för kvantitativa angreppssätt inom språkteknologi på en begreppslig nivå,

#### *Färdigheter och förmåga*

- tillämpa etablerade utvärderingsmetoder och tolka resultaten av dessa,
- tillämpa enkla statistiska språkteknologiska algoritmer och angreppssätt,
- använda verktyg för dataanalys och maskininlärning,
- använda kollaborativa verktyg för utveckling och dokumentation av mjukvara,
- utföra arbete enligt en på förhand bestämd tidplan,

#### *Värderingsförmåga och förhållningssätt*

- jämföra och välja lämpliga angreppssätt för maskininlärning och statistik.

### **Innehåll**

Kursen kommer att täcka följande områden:

- representation av data för maskininlärning och forskning om mänskligt språk,
- design av arbetsflöden (pipelines) för data,
- vetenskaplig metodologi inom maskininlärning för språkteknologi, med hänsyn till etiska och yrkesmässiga frågor,
- grunderna för maskininlärning i termer av informationsteori, sannolikhetslära, statistik och linjär algebra,
- klassiska maskininlärningsmetoder såsom Support Vector Machines (SVM) och logistisk regression,
- grunderna för perceptroner, neurala nätverk och Stochastic Gradient Descent.

### **Former för undervisning**

Undervisningen ges på distans eller på plats i form av föreläsningar, laborationer, inlämningsuppgifter, seminarier, projekt, övningar, eget arbete, eller arbete i grupp. De huvudsakliga undervisningsmetoderna är föreläsningar, demonstrationer och praktiska övningsuppgifter.

*Undervisningsspråk:* engelska

### **Former för bedömning**

Kursen examineras individuellt i skriftlig och/eller muntlig form.

Rättande lärare kan begära komplettering av examination.

Om student som underkänts två gånger på samma examinerande moment önskar byte av examinerare inför nästa examinationstillfälle, bör sådan begäran inlämnas skriftligt

till institutionen och ska bifallas om det inte finns särskilda skäl däremot (HF 6 kap 22§).

Om student fått rekommendation från Göteborgs universitet om särskilt pedagogiskt stöd kan examinator, i det fall det är förenligt med kursens mål och förutsatt att inte orimliga resurser krävs, besluta att ge studenten en anpassad examination eller alternativ examinationsform.

I det fall en kurs har upphört eller genomgått större förändringar ska student garanteras minst tre examinationstillfällen (inklusive ordinarie examinationstillfälle) under en tid av minst ett år, dock som längst två år efter det att kursen upphört/förändrats. Vad avser praktik och verksamhetsförlagd utbildning gäller motsvarande, men med begränsning till endast ett ytterligare examinationstillfälle.

Student förväntas lämna in efterfrågat material inom utsatt tid.

### **Betyg**

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

### **Kursvärdering**

Studenter som deltar i eller har avslutat kursen ges möjlighet att anonymt framföra erfarenheter av och synpunkter på denna i en kursvärdering. En sammanställning av kursvärderingen och kursansvarig lärares reflektion tillgängliggörs för studenterna inom rimlig tid efter kursslut. Nästa gång kursen ges presenteras sammanställningen och eventuella genomförda åtgärder.

### **Övrigt**

Kan ej ingå i examen tillsammans med LT2202, LT2212 eller LT2222.

Kursen kräver tillgång till dator (eller motsvarande) med internetuppkoppling.