

Introduction to Machine Learning

Fristående- och programkurs

9 hp

Introduction to Machine Learning

732A68

Gäller från: 2017 HT

Fastställd av

Filosofiska fakultetens kvalitetsnämnd

Fastställandedatum

2017-10-27

Huvudområde

Statistik

Utbildningsnivå

Avancerad nivå

Fördjupningsnivå

A1N

Förkunskapskrav

- 180 hp avslutade kurser varav 90 hp inom något av följande ämnen:
 - statistik
 - matematik
 - tillämpad matematik
 - datavetenskap
 - teknik
 - Godkända kurser i:
 - kalkyl
 - linjär algebra
 - Godkänd Statistik, grundkurs minst 6hp
 - Godkänd Programmering minst 6hp
 - Engelska 6
- Undantag ges för svenska

Lärandemål

Efter avslutad kurs skall den studerande på en avancerad nivå kunna:

- använda relevanta begrepp och metoder inom maskininlärning för att formulera, strukturera och lösa praktiska problem som har anknytning till stora eller komplexa datamängder,
- göra inferens för parametrarna i ett antal vanligt förekommande maskininlärningsmodeller,
- använda maskininlärningsmetoder för prediktion och beslutsstöd,
- utvärdera kvalitén av maskininlärningsmodeller,
- välja en lämplig modell i situationer med begränsad eller saknad information om bakomliggande samband i datamaterialet,
- implementera maskininlärningsmodeller i ett programmeringsspråk och använda existerande maskininlärningsmjukvaror för att analysera stora eller komplexa datamängder, göra prediktioner och utvärdera osäkerhet av dessa prediktioner.

Kursinnehåll

Kursen fokuserar på huvudbegrepp och huvudredskap i probabilistisk maskininlärning som är nödvändiga för ett professionellt arbete och forskning inom dataanalys.

- introduktion och översikt av maskininlärning (inklusive regression, klassificering, övervakad och oövervakad inlärning) och dess tillämpningsområden,
- Nearest Neighbors and Naïve Bayes,
- diskriminantanalys, logistisk regression och beslutsträd,
- modellval och osäkerhetsskattning: holdout metoden, korsvalidering, AIC, bootstrap och konfidensintervall,
- linjär regression och regulariseringsmetoder (Ridge, LASSO),
- splines, generaliserade linjära och additiva modeller,
- principalkomponentanalys (PCA) och principalkomponentregression (PCR),
- kernel utjämnare, kernel trick och supportvektormaskiner,
- neurala nätverk,
- bagging, boosting och random forests,
- Online learning och mixture models.

Undervisnings- och arbetsformer

Kursen består av föreläsningar, datorövningar och seminarier som kompletteras med självstudier. Föreläsningarna ägnas åt genomgång av teori, koncept och metodik. Datorövningarna ägnas åt praktisk dataanalys i en med hjälp av maskininlärningsmjukvara. Seminarier ägnas åt studentpresentationer och diskussioner av uppgifter.

Undervisningsspråk: engelska.

Examination

Skriftliga redovisning av labbuppgifter. Aktivt deltagande i seminarierna. En skriftlig tentamen. Detaljerad information återfinns i studiehandledningen.

Om LiU:s koordinator för studenter med funktionsnedsättning har beviljat en student rätt till anpassad examination vid salstentamen har studenten rätt till det. Om koordinatören istället har gett studenten en rekommendation om anpassad examination eller alternativ examinationsform, får examinator besluta om detta om examinator bedömer det möjligt utifrån kursens mål.

Studerande, vars examination underkänts två gånger på kursen eller del av kursen, har rätt att begära en annan examinator vid förnyat examinationstillfälle.

Den som godkänts i prov får ej delta i förnyat prov för högre betyg.

Betygsskala

ECTS, EC

Övrig information

Planering och genomförande av kurs ska utgå från kursplanens formuleringar. Den kursvärdering som ska ingå i varje kurs ska därför behandla frågan om hur kursen överensstämmer med kursplanen.

Kursen bedrivs på ett sådant sätt att både mäns och kvinnors erfarenhet och kunskaper synliggörs och utvecklas.

Institution

Institutionen för datavetenskap