



DATA- OCH INFORMATIONSTEKNIK

DIT865 Tillämpad maskininlärning, 7,5 högskolepoäng

Applied Machine Learning, 7.5 higher education credits

Avancerad nivå / Second Cycle

Fastställande

Kursplanen är fastställd av Institutionen för data- och informationsteknik 2017-01-05 att gälla från och med 2018-01-15, vårterminen 2018.

Utbildningsområde: Naturvetenskapligt 100 %

Ansvarig institution: Data- och informationsteknik

Inplacering

Kursen är obligatorisk inom masterprogrammet Applied Data Science. Den ges också som fristående kurs vid Göteborgs universitet.

Kursen kan ingå i följande program: 1) Software Engineering and Management masterprogram (N2SOF), 2) Computer Science, Master's Programme (N2COS) och 3) Applied Data Science masterprogram (N2ADS)

Huvudområde

Software Engineering

Data Science

Datavetenskap

Fördjupning

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

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

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

Förkunskapskrav

För tillträde till kursen krävs att studenten har en examen på kandidatnivå i ett ämne som är relevant för Data Science. Dessutom krävs

- minst 15 hp i programmering,

- en av kurserna DIT851 Introduktion till Data Science, 7,5 hp eller DIT855 Tillämpat matematisk tänkande, 7,5 hp, alternativt minst 7,5 hp i matematik,
- kursen DIT861 Statistiska metoder för Data Science, 7,5 hp alternativt minst 7,5 hp i sannolikhetsteori, statistik eller matematisk statistik.

Följande kunskapsnivå i Engelska krävs; Engelska 6/Engelska B eller motsvarande från ett erkänt internationellt test, t.ex. TOEFL, IELTS.

Lärandemål

Efter godkänd kurs ska studenten kunna:

Kunskap och förståelse

- beskriva de vanligaste typerna av maskininlärningsproblem,
- förklara vid vilken typ av tillämpningar som maskininläring kan användas, samt maskininläringens begränsningar,
- redogöra för hur maskininlärningsmetoder är beroende av informativ data och särdrag för att kunna tillämpas,
- förklara i stora drag hur olika maskininlärningsmodeller generaliserar utifrån träningsexempel.

Färdigheter och förmåga

- använda ett bibliotek för maskininläring i en applikationsdomän som är relevant för Data Science-området,
- skriva programkod för att implementera enstaka maskininlärningsalgoritmer,
- tillämpa utvärderingsmetoder för att bedöma kvaliteten hos maskininläringssystem, och jämföra flera olika sådana system.

Värderingsförmåga och förhållningssätt

- diskutera fördelarna och nackdelarna med olika maskininlärningsmodeller utifrån en given tillämpning,
- resonera om vilken typ av information eller särdrag som skulle kunna vara användbara i en maskininlärningsuppgift,
- välja en lämplig utvärderingsmetod för ett maskininlärningsbaserat system och motivera detta val,
- resonera om de etiska frågor som kan uppstå vid tillämpning av maskininlärningsbaserade system.

Innehåll

Kursen ger en inledning till tekniker och teorier inom maskininlärning, med ett fokus på dess praktiska tillämpningar. Under kursens gång kommer ett urval av ämnen att genomgå inom övervakad (supervised) inlärning, exempelvis linjär klassificering och regression, eller olinjära modeller som neurala nätverk, samt inom oövervakad (unsupervised) inlärning, till exempel klustringsmetoder. De typiska användningsfallen och begränsningarna hos dessa algoritmer kommer att diskuteras, och deras implementation kommer att undersökas med hjälp av programmeringsuppgifter. Metodologiska frågor som rör utvärdering av maskininlärningsbaserade system kommer också att diskuteras, samt några av de etiska frågor som kan dyka upp när man tillämpar den typen av teknologier.

Det kommer att vara ett starkt fokus på de användningsområden i verkligheten där maskininlärning kan tillämpas. Användningen av maskininlärningskomponenter i praktiska tillämpningar kommer att exemplifieras, och realistiska scenarion kommer att studeras i områden som t.ex. e-kommers, business intelligence, textanalys, bildanalys, eller bioinformatik. Vikten av konstruktion och urval av särdrag, samt deras pålitlighet, kommer att diskuteras.

Delkurser

1. **Skriftlig tentamen** (*Written examination*), 4 hp
Betygsskala: Väl godkänd (VG), Godkänd (G) och Underkänd (U)
2. **Skriftliga inlämningsuppgifter** (*Written assignments*), 3,5 hp
Betygsskala: Väl godkänd (VG), Godkänd (G) och Underkänd (U)

Former för undervisning

Föreläsningar, övningar, datorlaborationer.

Undervisningsspråk: engelska

Former för bedömning

Kursen examineras genom en individuell skriftlig salstentamen, samt obligatoriska skriftliga inlämningsuppgifter som redovisas genom rapportinlämning, varav vissa kommer genomföras individuellt och andra i grupper om normalt 2-4 studenter.

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 att få betyget Godkänd (G) för hela kursen måste studenten ha fått betyget G på båda delkurserna. För att få betyget Väl Godkänd (VG) för hela kursen krävs betyget VG på båda delkurserna.

Kursvärdering

Kursen utvärderas genom möten både under och efter kursen mellan lärare och studentrepresentanter. Därutöver används en anonym enkät för att få skriftlig information. Resultatet av utvärderingen används för att förbättra kursen genom att visa på delar som kan läggas till, förbättras, ändras eller tas bort.

Övrigt

Kurslitteratur kommer att publiceras senast 8 veckor innan kursstart.