

Kursplan

Data Science och maskininlärning 7,5 högskolepoäng, Grundnivå

Data Science & Machine Learning 7.5 Credits*, First Cycle

Lärandemål

Det övergripande målet är att studenten ska förvärva fördjupade kunskaper och färdigheter i att använda och utveckla mjukvara för data science samt grundläggande kunskaper inom data science, dvs ett tvärvetenskapligt förhållningssätt för att hitta, extrahera och upptäcka mönster i data genom användande av analysmetoder, domänkompetens och teknik.

Kunskap och förståelse

Efter avslutad kurs ska studenten kunna:

- förklara data science-livscykeln
- förklara Big Data och dataanalys
- förklara metoder för dataförberedelser
- beskriva metoder inom maskininlärning

Färdigheter och förmågor

Efter avslutad kurs ska studenten kunna:

- tillämpa icke-övervakade och övervakade maskinövervakningsalgoritmer för problemlösning
- tillämpa grundläggande begrepp i statistik och sannolikhetsteori, inklusive nyckelbegrepp som sannolikhetsfördelningar, statistisk signifikans, hypotesprovning och regression.
- använda programmeringsspråk för data science / dataanalys
- extrahera data från text
- använda utforskande dataanalys (EDA) för att beskriva data med hjälp av sammanfattningsstatistik och visualiseringstekniker

Värderingsförmåga och förhållningssätt

Efter avslutad kurs ska studenten kunna:

- tolka och analysera resultaten av en dataextraktionsprocess, samt utvärdera effekterna av val som gjorts under processen.

Innehåll

Kursen behandlar processer inom data science, dvs. en tvärvetenskaplig metod för att hitta, extrahera och upptäcka mönster i data genom en fusion av analytiska metoder, domenexpertis och teknik. I detta sammanhang omfattas områdena data mining, prognoser, maskininlärning, prediktiv analys, statistik och textanalys.

Inom ramen för den iterativa data-science-processen behandlas affärsförståelse med problemlösningsmetoder för specifikation av nyckelvariabler som ska fungera som modellmål samt identifiering av relevanta datakällor. Vidare ingår formulering av frågor som definierar affärs mål och kan kvantifieras av data-science tekniker.

För att kontrollera datakvalitet ingår förvärv av rådata, bearbetning av data (ETL), undersökning av data och modellering. För att underlätta utveckling av modell(er) och hitta den modell som bäst svarar på de initiala frågorna används sk feature engineering, då extraheras rådata och utmärkande särdrag skapas.

Avslutningsvis behandlas utvärdering av modellering och analys, presentation av resultat respektive driftsättning.

Examinationsformer

Inlämningsuppgifter (2,5 hp), laborationer (3 hp) och duggor (2 hp).

Arbetsformer

Föreläsningar, laborationer och inlämningsuppgifter.

Betyg

Som betygsskala används U - VG.

Slutbetygets nivå fastställs genom en sammanvägning av inlämningsuppgifter och duggor.

Förkunskapskrav

Objektorienterad programmering 7,5 hp grundnivå eller annan kurs i grundläggande

programmering
Statistisk analys 7,5 hp

Learning Outcomes

Ämnestillhörighet:

Informatik

Ämnesgrupp:

Informatik/Data- och systemvetenskap

Utbildningsområde:

Tekniska området, 100%

Kursen kan ingå i följande huvudområde(n):

1. Informatik
2. Mikrodataanalys

Fördjupningsbeteckning för respektive huvudområde:

1. G1F
2. G1F

Fastställd:

Fastställd 2020-09-24
Kursplanen gäller fr.o.m. 2020-11-11