



## INSTITUTIONEN FÖR FYSIK

### **FYM285 Bayesiansk dataanalys och maskininlärning, 7,5 högskolepoäng**

Learning from data, 7.5 credits

*Avancerad nivå / Second Cycle*

---

#### **Fastställande**

Kursplanen är fastställd av Institutionen för fysik 2019-03-11 och senast reviderad 2023-05-08. Den reviderade kursplanen gäller från och med 2024-01-15, vårterminen 2024.

*Utbildningsområde:* Naturvetenskapligt 100 %

*Ansvarig institution:* Institutionen för fysik

#### **Inplacering**

Kursen ingår i masterprogrammet i fysik.

Kursen kan ingå i följande program: 1) Physics, Master Program (N2PHY) och 2) Complex Adaptive Systems, Master Program (N2CAS)

#### *Huvudområde*

Fysik

#### *Fördjupning*

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

#### **Förkunskapskrav**

Kandidatexamen i fysik eller motsvarande, inklusive 30 hp matematik (inbegripet linjär algebra och analys) samt programmering.

Sökande måste påvisa kunskaper i engelska: engelska 6/engelska B från svenska gymnasiet, eller motsvarande nivå av ett internationellt erkänt test, till exempel TOEFL, IELTS.

#### **Lärandemål**

Kursen syftar till att introducera och lära studenterna att använda ett antal algoritmer och metoder inom statistisk inferens och maskininlärning, vilka i sin tur är centrala

inom vetenskaplig dataanalys. Stort fokus ligger på praktiska tillämpningar av Bayesiansk inferens inom ingenjör- och naturvetenskap, dvs förmågan att kunna kvantifiera styrkan hos induktiv inferens från fakta (t.ex. experimentell data) till propositioner såsom vetenskapliga hypoteser och modeller.

Kursen är projektbaserad och studenterna kommer att exponeras för modern naturvetenskaplig forskning genom olika projekt. Ett delsyfte är att lära studenterna att utveckla och strukturera datorprogram för att kunna genomföra pålitlig och vetenskaplig dataanalys. Mer specifikt kommer projektuppgifter att utföras med programmeringsspråket Python i kombination med användbara, och fritt tillgängliga, programbibliotek.

Efter godkänd kurs ska studenten kunna:

- utföra vetenskaplig dataanalys genom att utnyttja grundläggande koncept i Bayesiansk statistik tillsammans med kunskap kring vanligt förekommande statistiska fördelningar.
- kunna förklara centrala aspekter hos Monte Carlo-metoder och Markovkedjor, samt tillämpa dessa för att numeriskt sampla multivariata sannolikhetsfördelningar.
- kritiskt granska och kvantifiera osäkerheter hos modellparametrar från en statistisk analys av experimentell data; jämföra teoretiska modeller utifrån ett Bayesianskt synsätt.
- förstå och numeriskt implementera flera grundläggande algoritmer som används inom dataanalys och maskininlärning, som linjära regressions- och klassificeringsmetoder, enkla neurala nätverk och gaussprocesser.
- använda python för vetenskaplig dataanalys med statistiska inferensmetoder och för maskininlärning, samt för att visualisera resultat.
- skriva välstrukturerade tekniska rapporter där resultat och slutsatser från en vetenskaplig dataanalys presenteras klart och tydligt.
- vidhålla ett vetenskapligt och etiskt arbetssätt i processen att analysera data och att skriva datorprogram.

## Innehåll

Kursen består av två huvudsakliga delar:

1. Bayesiansk inferens och dataanalys
2. Maskininlärningsmetoder för dataanalys

Följande ämnen kommer att behandlas:

- Grundläggande koncept från statistik: förväntansvärden, varians, kovarians, korrelationsfunktioner och fel; diskreta och kontinuerliga sannolikhetsfördelningar;
- Genomgång av enkla statistiska modeller, binomialfördelning, Poissonfördelning, enkel och multivariat normalfördelning;

- Centrala delar av Bayesiansk statistik;
- Monte Carlo-metoder, Markovkedjor, Metropolis-Hastingsalgoritmen;
- Linjära regressions- och klassificeringsmetoder;
- Gauss- och Dirichletprocesser;
- Neurala nätverk.

### **Former för undervisning**

Föreläsningar.

Handledt arbete i datorsal (grupparbeten med numeriska projekt)

Ett antal mindre analytiska och numeriska hemuppgiftsproblem.

Två beräkningsprojekt med skriftlig rapportredovisning.

*Undervisningsspråk:* engelska

### **Former för bedömning**

Slutbetyget bestäms genom en kombination av de rättade hemuppgifterna och de skriftliga rapporterna.

Om student som underkänts två gånger på samma examinerande moment önskar byte av examinator inför nästa examinationstillfälle, ska sådan begäran inlämnas skriftligt till institutionen 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 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 VFU gäller motsvarande, men med begränsning till endast ett ytterligare examinationstillfälle.

### **Betyg**

På kursen ges något av betygen Väl godkänd (VG), Godkänd (G) och Underkänd (U). Slutbetyget bestäms genom en kombination av de rättade hemuppgifterna och de skriftliga rapporterna.

### **Kursvärdering**

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