

# Big Data Analytics

Programkurs

6 hp

Big Data Analytics

TDDE31

Gäller från: 2018 VT

**Fastställd av**

Programnämnden för data- och medieteknik,  
DM

**Fastställandedatum**

## Huvudområde

Informationsteknologi, Datateknik, Datavetenskap

## Utbildningsnivå

Avancerad nivå

## Fördjupningsnivå

AIX

## Kursen ges för

- Civilingenjör i datateknik
- Civilingenjör i informationsteknologi
- Civilingenjör i mjukvaruteknik
- Computer Science, masterprogram
- Datavetenskap, masterprogram

## Rekommenderade förkunskaper

Grundläggande databaskurs. Data mining eller maskininläring.

## Lärandemål

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

- samla och lagra Big Data i en distribuerad datormiljö
- genomföra grundläggande förfrågningar till en databas som opererar på ett distribuerat filsystem
- redovisa grundläggande principer för parallella beräkningar
- använda MapReduce begreppet för att parallellisera vanliga databearbetningsalgoritmer
- redovisa hur vanliga maskininlärningsmodeller bör modifieras för att bearbeta Big Data
- använda redskap för maskininläring av Big Data

## Kursinnehåll

Kursen fokuserar på huvudbegrepp och huvudredskap för lagring, bearbetning och analys av Big Data som är nödvändiga för ett professionellt arbete och forskning inom

dataanalys.

- Introduktion till Big Data: begrepp och verktyg
- Grundläggande principer av parallella beräkningar
- Filsystem och databaser för Big Data
- Förfrågningar för Big Data
- Resurshantering i en klustermiljö
- Parallella beräkningar för Big Data
- Maskininlärning för Big Data

## Undervisnings- och arbetsformer

Kursen består av föreläsningar och laborationer. Föreläsningarna behandlar teori, begrepp och metoder. Laborationerna ger praktisk erfarenhet i hantering av big data.

## Examination

TEN1	Skriftlig tentamen	U, 3, 4, 5	3 hp
LAB1	Laborationer	U, G	3 hp

## Betygsskala

Fyrgradig skala, LiU, U, 3, 4, 5

## Övrig information

Fortsättningskurser/kompletterande kurser: Svensk översättning saknas

## Institution

Institutionen för datavetenskap

## Studierektor eller motsvarande

Patrick Lambrix

## Examinator

Patrick Lambrix

## Undervisningstid

Preliminär schemalagd tid: 40 h  
Rekommenderad självstudietid: 120 h

## Kurslitteratur

### Övrigt

*Artikelsamling 2018.*