

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

A1X

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

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

Betygsskala

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

Övrig information

Fortsättningskurser/kompletterande kurser:

Avancerade datamodeller och databaser, parallell- och multicoreprogrammering

Om undervisningsspråk

Undervisningsspråk visas på respektive kurstillfälle på fliken "Översikt".

- Observera att även om undervisningsspråk är svenska kan delar av kursen ges på engelska.
- Om undervisningsspråk är Svenska/Engelska kan kursen i sin helhet ges på engelska vid behov.
- Om undervisningsspråk är Engelska ges kursen i sin helhet på engelska.

Övrigt

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

Planering och genomförande av kurs skall utgå från kursplanens formuleringar. Den kursvärdering som ingår i kursen skall därför genomföras med kursplanen som utgångspunkt.

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.