



LUNDS UNIVERSITET
Lunds Tekniska Högskola

Kursplan för

Programmeringsmodeller och metoder för big data

Programming Models and Practice for Big Data

TFRU20, 7,5 högskolepoäng, A (Avancerad nivå)

Gäller för: LTH:s uppdragsutbildningar 2023/24

Fakultet: Lunds tekniska högskola

Beslutad av: Utbildningsnämnd A

Beslutsdatum: 2022-11-09

Allmänna uppgifter

Fördjupning: Avancerad nivå, har endast kurs/er på grundnivå som förkunskapskrav.

Undervisningsspråk: Kursen ges på engelska

Syfte

I kursen lär sig deltagarna hur man analyserar och utformar program för big data problem. Studenterna får kunskap om arkitekturer, språk och ekosystem för big data, med fokus på systemet Spark. Tekniker har stor användning i en mängd olika områden såsom dataanalys, kundrekommendationer, trendförutsägelser, mönsterigenkänning, etc..

Mål

Kunskap och förståelse

För godkänd kurs skall studenten

visa kunskap om arkitekturer för big data behandling

visa förståelse för strukturen av big data programmeringsmodeller och ekosystem

Färdighet och förmåga

För godkänd kurs skall studenten

visa förmåga att hantera big data arkitekturer och utforma och skriva program för Spark

Värderingsförmåga och förhållningssätt

För godkänd kurs skall studenten

visa förmåga att välja ut och bedöma arkitekturer och algoritmer för big data problem

Kursinnehåll

Kursen består av fyra heldagsmöten som kommer att behandla:

1/ Cloudarkitekturer, begrepp i Spark och programmering i Spark

2/ Mer avancerade begrepp och programmering i Spark

3/ Övervakad maskininlärning med Spark: MLLib och MLLib-programmering

4/ Icke övervakad maskininlärning

Kursens examination

Betygsskala: UG - (U,G) - (Underkänd, Godkänd)

Prestationsbedömning: Inlämningsuppgifter i form av program och rapporter. 4 inlämningsuppgifter är av programmeringstyp som utföres enskilt eller i par. 4 inlämningsuppgifter är av rapportskrivningstyp som görs individuellt.

Om så krävs för att en student med varaktig funktionsnedsättning ska ges ett likvärdigt examinationsalternativ jämfört med en student utan funktionsnedsättning, så kan examinator efter samråd med universitetets avdelning för pedagogiskt stöd fatta beslut om alternativ examinationsform för berörd student.

Antagningsuppgifter

Förutsatta förkunskaper: Goda kunskaper i objekt-orienterad programmering med vana vid Java, Scala, eller Python. Grundläggande kunskaper om statistik. Till exempel är de förutsatta förkunskaperna uppfyllda genom följande LTH-kurser: - Programmeringsteknik för C, D (EDA016) - Programmeringsteknik - fördjupningskurs (EDAA01) - Matematisk statistik, allmän kurs (FMS012)

Kurslitteratur

- Sandy Ryza, Uri Laserson, Sean Owen, Josh Wills: Advanced Analytics with Spark, Patterns for Learning from Data at Scale. O'Reilly Media, 2015, ISBN: 978-1-491-91276-8.
- Holden Karau, Andy Konwinski, Patrick Wendell, Matei Zaharia: Learning Spark, Lightning-Fast Big Data Analysis. O'Reilly Media, 2015, ISBN: 978-1-449-35862-4.

Kontaktinfo och övrigt

Kursansvarig: Pierre Nugues, pierre.nugues@cs.lth.se