



INSTITUTIONEN FÖR DATA- OCH INFORMATIONSTEKNIK

DIT066 Beräkningsmetoder för storskaliga data, 7,5 högskolepoäng

Computational techniques for large-scale data, 7.5 credits

Avancerad nivå / Second Cycle

Fastställande

Kursplanen är fastställd av Institutionen för data- och informationsteknik 2023-11-14 att gälla från och med 2025-01-20, vårterminen 2025.

Utbildningsområde: Naturvetenskapligt 100 %

Ansvarig institution: Institutionen för data- och informationsteknik

Inplacering

Kursen ingår i flera program och ges som fristående kurs vid Göteborgs universitet.

Kursen kan ingå i följande program: 1) Computer Science, Master's Programme (N2COS), 2) Software Engineering and Management masterprogram (N2SOF) och 3) Applied Data Science masterprogram (N2ADS)

Huvudområde

Data Science

Datavetenskap

Fördjupning

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

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

Förkunskapskrav

För tillträde till kursen krävs att studenten har en examen på kandidatnivå inom något ämne, eller har minst 90 hp i datavetenskap, software engineering eller motsvarande. Specifikt krävs minst 15 hp avklarade kurser i programmering, varav minst 7.5 poäng i programmering python eller motsvarande. Därutöver krävs även att studenten har klarat en kurs i sannolikhetsteori eller statistik, exempelvis DIT862 eller MSG810 eller motsvarande.

Följande kunskapsnivå i Engelska krävs; Engelska 6/Engelska B eller motsvarande från ett erkänt internationellt test, t.ex. TOELF, IELTS.

Lärandemål

Efter godkänd kurs ska studenten kunna:

Kunskap och förståelse

- diskutera de viktigaste teknologiska aspekterna vid design och implementering av system för analys av storskaliga data,
- förklara skillnader mellan parallella programmeringsmodeller
- beskriva datastrukturer och algoritmer för storskaliga data och deras användning.

Färdigheter och förmåga

- implementera tillämpningar för att transformera och analysera storskaliga data med hjälp av olika parallella mjukvaruramverk,
- kunna använda algoritmer och datastrukturer för beräkningar med storskaliga data.

Värderingsförmåga och förhållningssätt

- föreslå lämpliga beräkningsinfrastruktur och metoder för analys och diskutera deras fördelar och nackdelar,
- diskutera fördelar och nackdelar av olika strategier för parallelisering,
- kunna välja mellan olika algoritmer och metoder baserad på parallelisering i syfte med att accelerera beräkningsmässiga workloads.

Innehåll

Kursens mål är att fördjupa studenternas kunskaper och färdigheter inom den tekniska sidan av data science, inklusive de relevanta datamodellerna, samt lämpliga mjukvaru- och hårdvarumiljöer. Kursen kommer att introducera aspekter av design och implementation av storskaliga data science-lösningar.

Framför allt kommer kursen att innefatta:

- en överblick av datorarkitekturer, algoritmiska tillvägagångssätt, och högpresterande beräkningsinfrastruktur, med ett fokus på de begränsningar som finns vid behandling av storskaliga data,
- en introduktion i relevanta ramverk för klusterberäkning med storskaliga data,
- implementering av dataanalysverktyg på ett kluster med hjälp av Python och lämpliga mjukvaruramverk,
- datastrukturer och algoritmer som exempelvis index strukturer, som avsevärt kan accelerera beräkningar för storskaliga data.

Delkurser

1. Skriftlig salstentamen (Written hall examination), 3 hp

Betygsskala: Mycket väl godkänd (5), Väl godkänd (4), Godkänd (3) och Underkänd (U)

2. Inlämningsuppgifter (*Assignments*), 4,5 hp

Betygsskala: Mycket väl godkänd (5), Väl godkänd (4), Godkänd (3) och Underkänd (U)

Former för undervisning

Föreläsningar, datorlaborationer och övningar.

Undervisningsspråk: engelska

Former för bedömning

Kursen examineras genom en skriftlig salstenta och genom obligatoriska skriftliga inlämningsuppgifter. Några av inlämningsuppgifterna utförs individuellt och andra utförs i grupper av 2-4 studenter. Det kommer att finnas icke-obligatoriska inlämningsuppgifter som ger bonuspoäng för den skriftliga tentan. Bonuspoängen gäller för de två tentorna och omtentorna som är schemalagda närmast efter kurstillfället där bonuspoängen förvärvades.

Om student som underkänts två gånger på samma examinerande moment önskar byte av examinator inför nästa examinationstillfälle, bör sådan begäran inlämnas skriftligt till institutionen och ska bifallas om det inte finns särskilda skäl däremot (HF 6 kap 22§).

Om student fått rekommendation från Göteborgs universitet om särskilt pedagogiskt stöd kan examinator, i det fall det är förenligt med kursens mål och förutsatt att inte orimliga resurser krävs, besluta att ge studenten en anpassad examination eller alternativ examinationsform.

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 verksamhetsförlagd utbildning gäller motsvarande, men med begränsning till endast ett ytterligare examinationstillfälle.

Betyg

På kursen ges något av betygen Mycket väl godkänd (5), Väl godkänd (4), Godkänd (3) och Underkänd (U).

För att få godkänt på kursen måste samtliga obligatoriska moment vara godkända. För att få ett högre betyg än godkänt så krävs ett högre viktat genomsnitt på de obligatoriska momenten.

Kursvärdering

Kursen utvärderas genom möten, både under och efter kursen, mellan lärare och studentrepresentanter. Ett anonymt skriftligt frågeformulär skickas även ut till studenterna efter kursens slut. Resultaten av utvärderingarna används för att förbättra kursinnehållet och som indikation till vilka delar som skulle kunna läggas till, tas bort, förbättras eller ändras.

Övrigt

Kursen är samläst med Chalmers.

Kurslitteratur kommer att publiceras senast 8 veckor innan kursstart.

Kursen ersätter kursen DIT065, 7,5 hp. Den här kursen kan inte ingå i en examen som innehåller DIT065. Den kan inte heller ingå i en examen som bygger på en annan examen där DIT065 ingår.