



## Dataanalysprojekt

### Data analysis project

7,5 högskolepoäng

7,5 credits

---

**Ladokkod:** C1DP1B

**Version:** 1.1

**Fastställt av:** Utskottet för utbildningar inom bibliotek, information och IT 2020-12-10

**Gäller från:** VT 2021

**Nivå:** Grundnivå

**Huvudområde (successiv fördjupning):** Datalogi (G2F), Informatik (G2F)

**Utbildningsområde:** Naturvetenskap

**Ämnesgrupp:** Informatik/Data- och systemvetenskap

**Förkunskapskrav:** Avklarade kurser om minst 60 högskolepoäng inom informatik som inkluderar Maskininläring 7,5 högskolepoäng, Datorgrafik och Visualisering 7,5 högskolepoäng, samt Informationssystem och data 7,5 högskolepoäng.

**Betygsskala:** Underkänd, Godkänd eller Väl godkänd

---

### Innehåll

Kursen ger en praktisk fördjupning inom informationsutvinning, med utgångspunkt i tidigare kunskaper i maskininläring och datavisualisering. Kursen introducerar ytterligare kunskaper inom icke-väglett lärande (eng. unsupervised learning) med fokus på tillämpning och utvärdering av klusteranalys och associationsregelanalys. För praktisk tillämpning introduceras ett för kursen valt ramverk för automatiserade processer och arbetsflöden som inkluderar funktionalitet för maskininläring och datavisualisering.

Praktiska färdigheter utvecklas genom ett större dataanalysprojekt, där studenter i grupp analyserar en större datamängd med fokus på att utvinna information genom att kombinera olika tekniker och metoder från maskininlärnings- och dataanalysområdet.

### Mål

Efter avklarad kurs ska studenten kunna, med avseende på:

#### *Kunskap och förståelse*

- 1.1. redogöra för grundläggande teorier och begrepp rörande informationsutvinning.
- 1.2. redogöra för grundläggande teorier och begrepp rörande datapreparering och datarensning.
- 1.3. redogöra för grundläggande teorier och begrepp rörande klusteranalys och dess utvärdering.
- 1.4. redogöra för grundläggande teorier och begrepp rörande associationsregelanalys och dess utvärdering.

#### *Färdighet och förmåga*

- 2.1. tillämpa algoritmer, tekniker och metoder för datapreparering och datarensning.
- 2.2. tillämpa algoritmer, tekniker och metoder för klusteranalys.
- 2.3. tillämpa algoritmer, tekniker och metoder för associationsregelanalys.
- 2.4. tillämpa algoritmer, tekniker och metoder för klassificering och regressionsanalys.
- 2.5. tillämpa algoritmer, tekniker och metoder för datavisualisering.
- 2.6. använda etablerade ramverk för automatiserade processer för databehandling och maskininläring.

#### *Värderingsförmåga och förhållningssätt*

- 3.1 välja och utnyttja lämpliga metoder för förbehandling och utvärdering givet en datamängd.
- 3.2 välja och utnyttja lämpliga algoritmer och tekniker för informationsutvinning givet en datamängd.

### Undervisningsformer

Undervisningen på kursen består av:

- föreläsningar
- handledning
- seminarium

Föreläsningarna presenterar kursens teoretiska innehåll. Handledning sker för projektet, vilket utgör kursens praktiska innehåll. Det genomförda projektet redovisas i detalj skriftligen och presenteras även översiktligt på ett seminarium, muntligt och skriftligt.

Undervisningen bedrivs på svenska, men undervisning på engelska kan förekomma.

### **Examinationsformer**

Kursen examineras genom följande moment:

Tentamen: skriftlig individuell tentamen

Lärandemål: 1.1 - 1.4, 2.1 – 2.5

Högskolepoäng: 2,5

Betygsskala: UVG

Rapport: skriftlig redovisning av dataanalysprojekt i KNIME (gruppuppgift)

Lärandemål: 2.1 – 2.6, 3.1 – 3.2

Högskolepoäng: 4

Betygsskala: UG

Seminarium: muntlig och skriftlig presentation av dataanalysprojekt i KNIME (gruppuppgift)

Lärandemål: 2.1 – 2.6, 3.1 – 3.2

Högskolepoäng: 1

Betygsskala: UG

För betyget Godkänd på hel kurs krävs Godkänd på samtliga examinationsmoment. För betyget Väl Godkänd på hel kurs krävs dessutom Väl Godkänd på tentamen.

Om studenten har ett beslut/rekommendation om särskilt pedagogiskt stöd från Högskolan i Borås på grund av funktionsnedsättning, har examinator rätt att anpassa examinationen. Examinator har att utifrån kursplanens mål avgöra om examinationen kan anpassas i enlighet med beslutet/rekommendationen.

Studentens rättigheter och skyldigheter vid examination är enligt riktlinjer och regelverk vid Högskolan i Borås.

### **Kurslitteratur och övriga läromedel**

Kurslitteraturen är på engelska.

Greco, C. (2020). Data Science Tools: R • Excel • KNIME • OpenOffice, Mercury Learning & Information. (206 sidor)

Tan, P.-N., Steinbach, M., Kumar, V. & Karpatne, A. (2019). Introduction to Data Mining, Global Edition, Pearson. (800 sidor)

Vetenskapliga artiklar och föreläsningmaterial kan tillkomma enligt lärarens anvisningar, max 200 sidor.

### **Studentinflytande och utvärdering**

Kursen utvärderas i enlighet med gällande riktlinjer för kursvärderingar vid Högskolan i Borås, där studenternas synpunkter ska inhämtas. Kursutvärderingsrapporten publiceras och återkopplas till deltagande och blivande studenter i enlighet med ovan nämnda riktlinjer, och ligger till grund för framtida utveckling av kurser och utbildningsprogram. Kursansvarig lärare ansvarar för att utvärdering enligt ovan genomförs.

### **Övrigt**

Kursen ingår i Systemarkitekturutbildningen.