

# Studieplan for Big Data and Analytics (15 studiepoeng) (Høst 2017)

## Fakta om programmet

**Studiepoeng:** 15

**Studiets varighet:** ½ år

**Undervisningsspråk:** Engelsk

**Studiested:** Høgskolen i Østfold, studiested Halden

## Kontakt

Antonella Serrano

**Telefon:** +47 696 08 096

**E-post:** videre@hiof.no

### *Studieprogramansvarlig*

HiØ VIDERE

## Innholdsfortegnelse

- Informasjon om studiet
- Hva lærer du?
- Opptak
- Oppbygging og gjennomføring
- Jobb og videre studier
- Studieplanen er godkjent og revidert
- Studiemodell

## Informasjon om studiet

Dette er et betalingsstudium. For nærmere informasjon om studieavgift, vennligst se [studiekatalog](#) eller kontakt [HiØ VIDERE](#).

Samfunnet er og blir i større og større grad digitalisert. Flere og flere av våre aktiviteter foregår på digitale arenaer. Fra kontakt med venner, familie og kolleger gjennom sosiale medier, til levering av selvangivelse, registrering av strømforbruk og bruk av medisinske målesystemer. Behovet for analyse av store datamengder blir bare større og større i dagens samfunn.

Målgruppen er personer som ønsker å øke sin kompetanse innen metoder for analyse av store datamengder.

Anbefalte forkunnskaper ut over opptakskravet (se nedenfor) er grunnleggende kompetanse i:

- programmering,
- lineær algebra,
- sannsynlighetsteori,
- statistikk,
- kalkulus og
- optimalisering.

## Hva lærer du?

### Studiets læringsutbytte

#### Kunnskap

Studenten

- har oversikt over alle de viktigste metodene innen dataanalyse.
- har kunnskap om de utfordringer som i praksis finnes innen dataanalyse, for eksempel overfitting, manglende verdier og klassifiseringer som har forskjellig kostnad.
- har fått bedre kunnskaper i grunnleggende emner som numerisk optimering og statistiske metoder relatert til dataanalyse, for eksempel Bayesianske metoder brukt av søkemotorer som Google og i spam filter.

#### Ferdigheter

Studenten kan

- tilrettelegge eller kode om data slik at det passer for dataanalyse og maskinlæringsalgoritmer.
- velge rett verktøy for en gitt type data.
- avgjøre hvor gode resultater man faktisk har ved hjelp av enkel statistisk analyse av for eksempel klassifiseringsnøyaktighet.
- bruke dataanalyse i praktiske applikasjoner.

#### Generell kompetanse

Studenten har kompetanse til å behandle og analysere data av vilkårlig type.

## Opptak

Generell studiekompetanse eller realkompetanse.

## Oppbygging og gjennomføring

### Studiets oppbygging og innhold

Studiet består av ett emne med et omfang på 15 studiepoeng.

Dette emnet gir en bred introduksjon i dataanalyse og statistisk mønstergjenkjenning. Du blir kjent med klassiske, så vel som moderne dataanalysemetoder. I tillegg vil du lære om metodenes fordeler og ulemper, samt om når og hvordan du skal benytte dem.

Emnet tar for seg følgende temaer:

Del 1: Course Introduction, linear regression and logistic regression

Del 2: Overfitting problem, statistical view and probability density estimation

Del 3: Decision tree, Random forest, Bayesian Network and Hidden Markov Model

Del 4: Data pre-processing, generalization and advanced models

## Organisering og læringsformer

Studiet går over ett semester med fire to-dagers samlinger på dagtid, totalt 8 dager. Det gis veiledning via e-post mellom samlingene.

**Arbeidskrav:** Minimum 75% deltakelse på undervisningssamlinger. Tilstedeværelse vil bli registrert. For at studenten skal kunne framstille seg til eksamen må arbeidskravet være godkjent senest 3 uker før eksamen.

## Evaluering av studiet

For å kunne tilby en aktuell og relevant utdanning av god kvalitet er vi avhengige av studentenes tilbakemeldinger. Evaluering gjennomføres i henhold til høgskolens kvalitetssystem. Etter fullført studium evaluerer studentene studiet på elektronisk skjema fra HiØ VIDERE.

## Vurdering

### Mappevurdering

Mappen består av fire prosjekter. Det første prosjektet vil bli gitt etter første samling, det andre etter andre samling osv. Prosjektene skal gjennomføres individuelt eller i grupper på to og to.

Mappen vurderes som en helhet, og det gis én samlet, individuell karakter for mappen.

Karakterregel: Bestått/Ikke bestått.

Det benyttes enten to interne sensorer, eller en ekstern og en intern sensor.

Ved nyeksamen må hele mappen gjennomføres på nytt.

### Plagiatkontroll

Alle skriftlige arbeidskrav og eksamensoppgaver kan plagiatkontrolleres. Plagiering og avskrift av faglitteratur og andre skriftlige arbeider uten korrekt bruk av referanser/kilder vil bli vurdert som forsøk på fusk. Se for øvrig [Forskrift om eksamen og studierett ved Høgskolen i Østfold](#)

## Litteratur

Studiet baserer seg i hovedsak på forelesningsmateriale og webressurser.

Anbefalt litteratur:

- Tom. M. Mitchell, Machine Learning, 1997, McGraw-Hill Higher Education; ISBN: 0070428077

- M. Bishop, Neural Networks for Pattern Recognition, 1996.

- H. Witten, E. Frank and M. A. Hall, Data Mining: Practical Machine Learning Tools and Techniques, 2011.

- A. Ng, Machine Learning, Stanford University, online course.

- D. Koller, Probabilistic Graphical Models, Stanford University, online course.

- G. Hinton, Neural Networks for Machine Learning, University of Toronto, online course.

- M. Collins, Neural language Processing, Columbia University, online course.

## Jobb og videre studier

Ønsker du ytterligere kompetanse innen dataanalyse, kan du søke deg til [bachelorstudium i informatikk - design og utvikling av IT-systemer](#) eller [masterstudium i Applied Computer Science ved Høgskolen i Østfold](#) dersom du fyller gjeldende krav til opptak.

Personer med kompetanse innenfor analyse av store datamengder vil i en økende digitalisert verden bli svært ettertraktede ressurser.

## Studieplanen er godkjent og revidert

### Studieplanen er godkjent

Dekan Harald Holone, 22.03.16

### Studieplanen er revidert

Studieleder Monica Lind Kristensen, 24.04.17

### Studieplanen gjelder for

Høsten 2017

## Studiemodell

Denne studiemodellen har en ny utforming. [Fortell oss hva du synes om den](#)

Høst 2017

Obl. emne

ITVBIGD116  
Big Data Analytics

15 stp

# ITVBIGD116 Big Data Analytics (Høst 2017)

## Fakta om emnet

**Studiepoeng:** 15

**Ansvarlig avdeling:** Avdeling for informasjonsteknologi

**Undervisningsspråk:** Norsk

**Varighet:** ½ år

## Innholdsfortegnelse

- Innhold

## Innhold

Studiet *Big Data Analytics* (15 studiepoeng) består av ett emne. For nærmere beskrivelse, se studieplan.