

SOS2901 – Anvendt maskinlæring for samfunnsvitere

Maskinlæring er sentralt i å nyttiggjøre seg store data, deriblant for automatisering generelt og datadrevne beslutningssystemer både i privat og offentlig sektor. En viktig anvendelse er klassifisering og prediksjoner for nye observasjoner der man ennå ikke vet utfallet. Noen ganger vil det innebære å spå om fremtiden og gjøre beslutninger basert på prediksjoner.

Med økt digitalisering vil innsikt i denne type dataanalyse være viktig for samfunnsforståelse generelt. Det er også metoder med økende utbredelse i samfunnsvitenskapelig forskning. Selv om samfunnsvitere i liten grad vil stå for utvikling av slike systemer i arbeidslivet, men vil kunne være involvert i anbudsrunder, implementering eller andre vurderinger av slike systemer. Samfunnsvitere vil derfor ha stor praktisk nytte av innsikt i maskinlæring.

Dette kurset gir en praktisk innføring i grunnleggende prinsipper for maskinlæring og håndtere noen relativt enkle algoritmer for strukturerte data. Det vektlegges vurderinger av akseptable feilrater (falske positive vs. falske negative), og systematiske forskjeller i presisjon på tvers av undergrupper (bias og fairness). Gjennomgående vektlegges justering av algoritmene (tuning) for å oppnå akseptable feilrater og veie ulike hensyn mot hverandre.

For noen anvendelser, som f.eks. målrettet reklame, er det lite alvorlig om prediksjonene er feil. I andre typer anvendelser skal det derimot tas beslutninger som potensielt innebærer alvorlige konsekvenser for enkeltindivider, f.eks. kredittvurderinger, rekruttering i arbeidslivet, eller risikovurderinger for fremtidig kriminalitet. Algoritmene må derfor vurderes i lys av hva prediksjonene skal brukes til, og hvilke konsekvenser det får. Det gjelder særlig konsekvenser hvis prediksjonene er feil. Slike vurderinger gjelder prinsipielt sett også for andre beslutningssystemer, herunder skjønn. Selv når feilratene er store er det ikke alltid åpenbart at alternativene til maskinlæring er bedre. Slike vurderinger står sentralt i emnet.

Kurset vektlegger praktiske anvendelser med håndtering av datasett og bruk av softwaren *R*. Det tekniske nivået er moderat. Undervisningen tar utgangspunkt i at studentene er kjent med grunnleggende lineær regresjon, og er kjent med *R*. Det anbefales derfor å ha tatt SOSGEO1120 eller tilsvarende (f.eks. STV1020).

Læringsmål

Etter endt kurs skal studentene kunne følgende:

- Kjenne viktige prinsipper for maskinlæring, herunder forskjell på supervised og unsupervised ML, forstå bias-variance trade-off, overfitting osv.
- Beherske grunnleggende modeller for prediksjon og klassifikasjon, med vekt på regresjon, tre-baserte algoritmer og boosting.
- Kjenne grunnleggende teknikker for tolkbar maskinlæring: feature importance og partial dependence
- Kunne vurdere prediksjoners pålitelighet, med vekt på feilrater og mål på algorithmic fairness. Det vektlegges tuning av algoritmene for å oppnå ønsket resultat.
- Beherske grunnleggende teknikker for klustering og datareduksjon
- Forstå viktige muligheter og begrensninger ved datadrevne modeller. Herunder datakvalitet, bias og utilsiktede konsekvenser.

Undervisning

Undervisningen vil bestå av undervisningsvideoer og seminarer som er fokusert på praktisk oppgaveløsning der det jobbes med oppgaver og R-programmering. Seminarene er ikke obligatoriske, men deltagelse krever innlevering av en enkel oppgave i forkant av seminaret. Seminaret forutsetter aktiv deltagelse fra alle og det er ikke mulig å følge seminaret uten forberedelse.

Forutsetninger

Det forutsettes en grunnleggende kjennskap til kvantitative metoder og statistikk. Studentene bør ha noe kjennskap til programmet R. Det forventede nivået tilsvarer fullført SOSGEO1120 eller STV1020. Studenter som ikke har tatt disse emnene bør sjekke læringsmålene på emnesiden til SOSGEO1120.

Studentene må ha tilgang til egen datamaskin med internettilgang med R og Rstudio installert.

Eksamen

4-timers skoleeksamen.

Eksamen vil inkludere flervalgsoppgaver, korte tekstsvar og analyse av datasett i R.