

Kursplan

Statistisk- och maskininlärning 10 högskolepoäng, Forskarnivå

Statistical and Machine Learning 10 Credits, Third Cycle

Lärandemål

Efter avslutad kurs ska doktoranden kunna:

- Välja lämpliga statistiska modeller och metoder för dataanalys av praktiska problem baserat på välgrundade argument, särskilt när den underliggande datagenere-rande mekanismen är okänd.
- Använda olika statistiska inlärningsalgoritmer (både övervakat och oövervakat lä-rande) på praktiska problem.
- Utvärdera och optimera inlärningsmodellernas och algoritmernas prestanda samt kommunicera modellens/algoritmens förväntade osäkerhet.
- Kombinera olika modeller för att uppnå högre prediktiv noggrannhet.
- Tillämpa neurala nätverk för problemlösning i reella fall.
- Göra jämförande analys, både teoretiskt och empiriskt, för att avgöra vilket neuralt nätverk som är bäst lämpat för ett visst problem.
- Designa olika typer av neurala nätverk, utvärdera deras prestanda och använda dem för att lösa komplexa problem.
- Använda deep learning för problemlösning i reella fall.

Innehåll

Kursen består av två delar. Den första delen är statistiskinlärning som omfattar lärandemålen 1-4 och den andra delen är maskininlärning som omfattar lärandemålen 5-8. Båda delarna har fokus på tillämpning av statistiskinlärning respektive maskininlärning där den senare betonar neurala nätverk och deep learning.

Del ett fokuserar främst på de tillämpade aspekterna av statistiskinlärning men algoritmerna för statistiskinlärning studeras också. Den omfattar övervakade (supervised) inlärningsalgoritmer, med särskild tonvikt på klassificeringsmetoder som logistisk regression, klassificerings-träd, linjär diskriminantanalys, kvadratisk diskriminantanalys, K nearest neighbour, support vector machine, och regressionsmetoder innefattande linjär regression, splines, generaliserade additiva modeller och regressionsträd. Den behandlar också oövervakade (unsupervised) inlärningsmetoder som principalkomponentanalys, k mean klustering och hierarkisk klustering. Modellvalidering genom korsvalidering, och bootstrap metoder behandlas. Regularisering för modellval, högdimensionell dataanalys, och förbättring av prediktionsprestanda genom model averaging, bagging och boosting

ingår också.

Del två ger en introduktion till maskininlärning och en översikt över neurala nätverk. Perceptronet som grundelement för linjär separabilitet och dess begränsningar i klassificering diskuteras. Sedan studeras olika aktiveringsfunktioner och sigmoid perceptron för att lösa icke-linjära klassificeringsproblem. Olika typer av maskininlärningsparadigmer som övervakad (supervised), oövervakad (unsupervised) och reinforcement inlärande omfattas. Feedforward neurala nätverk och algoritmen för backpropagation kommer att presenteras. Kursen kommer också att omfatta återkommande neurala nätverk (RNN). Slutligen diskuteras deep learning med tonvikt på de grundläggande principerna och olika typer av neurala nätverk för deep learning.

Examinationsformer

Delarna 1 och 2 examineras var för sig.

Del 1 examineras med inlämningsuppgift och skriftlig tentamen om 5 hp.

Del 2 examineras med projektarbete, skriftlig reflektion och seminarium om 5 hp.

Arbetsformer

Kursen består av föreläsningar, laborationer, projektarbete och seminarier.

Betyg

Som betygsskala används U–G.

Förkunskapskrav

Magisterexamen i Mikrodataanalys

Ämnestillhörighet:

Mikrodataanalys

Fastställd:

Fastställd 2019-10-29

Kursplanen gäller fr.o.m. 2019-10-29