



Kursus

1 dag

Nr. 38008 P

DKK 4.150

ekskl. moms

Dato

23-04-2021

05-10-2021

Sted

Taastrup

Skejby

R3 - Machine learning og predictive models i R

Kurset er tredje og sidste kursus af i alt tre statistikkurser i R. Store mængder data medfører udfordringer. Kurset er til dig, der ønsker at komme i gang med din Big Data analyse. Du lærer at bruge R til at modellere data i to populære machine learning algoritmer: Random Forest og Boosting. Efter ønske kan kurset følges online som et webinar på dagen.

Udbytte

Du har lært, at machine learning dækker over algoritmer udviklet til prædiktion og data mining særligt i store datamængder. Du har lært om the curse of dimensionality og the variance-bias-trade-off, som grundlæggende problemer i enhver data analyse og særligt i en data analyse af store datamængder. Du har lært rationale i krydsvalideringsproceduren og lært at gennemføre en krydsvalidering i R. Du har lært populære machine learning algoritmer at kende og anvende i R: Random Forest og Boosting.

Forudsætninger

Du skal medbringe egen PC. Du skal have kendskab til R og en forståelse af statistiske grundprincipper svarende til niveauet i kurset R1 - Introduktion til statistikprogrammet R og R2 - Introduktion til lineære modeller i R.

Program

Programmet består af flere oplæg fra underviseren og en række Hands-on forløb bestående af et udleveret R-script, som deltageren kører igennem under underviserens vejledning.

10:00 Velkomst

10:15 Big data og machine learning - en ny æra begynder: Nye algoritmer til nye udfordringer, Machine learning og predictive modelling, Krydsvalidering, Et eksempel

11:00 Pause

11:10 Hands-on 1: Krydsvalidering i en lineær model i R

12:00 Frokost

12:30 Machine learning 1: Random Forest, Random Forest algoritmen, Et eksempel

13:00 Hands-On 2: Random Forest model i R

13:50 Pause

14:00 Machine learning 2: Boosting, Boosting algoritmen, Et eksempel

14:50 Pause

15:00 Hands-On 3: Boosting i R

15:50 Afrunding

16:00 Slut

Frokost

Frokostbuffet er inkluderet i prisen.

Min. antal deltager

8. Teknologisk Institut forbeholder sig retten til at aflyse kurset såfremt der er færre end det minimale antal deltager.

Foreslået læsning

Gareth, J., Witten, D., Hastie, T., Tibshirani, R. 2013. An Introduction to Statistical Learning with Applications in R. Springer, NY. Gratis online version: <http://www-bcf.usc.edu/~gareth/ISL/>

Har du faglige spørgsmål så kontakt



Philipp Trénel
+45 72203392
phtr@teknologisk.dk