

## PS5. Data Science in the Wild

Organised by Cecilie Utke Rank (Rigshospitalet/University of Copenhagen), Federica Belmonte (Danish Cancer Research Center) and Sebastian Weichwald (University of Copenhagen)

### Session overview

Scalability, missing values, small datasets, inconsistent data records, sensitive data, unreliable black-box models versus reliable yet more restrictive models, computational expense, cloud infrastructure, ... — This session is for knowledge sharing about the hurdles, problems, and surprises we face when moving from theory and methodology developed on clean toy datasets to deploying data science solutions in the wild. How is data science used in practice? What issues are predominant? What are the lessons learnt and advice for fellow data scientists?

### Speakers

- Monica Jane Emerson, Novo Nordisk
- Ulrik Kristoffer Stoltze, Rigshospitalet
- Mads Schreiner Simonsen, ATP
- Federica Belmonte, MedEngine
- Kristoffer Hougaard Madsen, DTU CogSys & DRCMR

### Programme

- 13:45 Kick-off
- 14:00 Monica Jane Emerson, Novo Nordisk
- 14:25 Ulrik Kristoffer Stoltze, Rigshospitalet
- 14:50 Mads Schreiner Simonsen, ATP
- 15:15 Federica Belmonte, MedEngine
- 15:30 Kristoffer Hougaard Madsen, DTU CogSys & DRCMR
- 15:55 Wrap-up