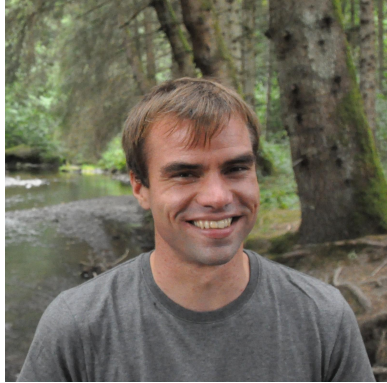**TWO-DAY COURSE, Saturday and Sunday**
Fundamentals of Machine Learning for Mass Spectrometry Data Analysis

Instructors

Wout Bittremieux
University of Antwerp

Will Fondrie
Talus Bioscience

Over the past decade, machine learning has become a dominant technology for data-intensive discovery in nearly all scientific domains. Today, almost all biomedical research employs machine learning techniques to derive new knowledge from complex biological data. This course will introduce the fundamentals of common machine learning techniques used for the analysis of mass spectrometry data. The main goal of the course is to promote basic data literacy for people new to machine learning. This will help researchers to perform basic machine learning analyses, recognize common pitfalls, know when and how to consult a machine learning expert, and better understand machine learning applications in the scientific literature.

The course sessions will cover fundamental methods in machine learning applied to proteomics and metabolomics examples. Each session will include hands-on exercises to illustrate the covered material. Topics that will be covered include:

- Introduction
    - What is machine learning?
    - Classification vs regression
    - Bias–variance trade-off

- Unsupervised machine learning
    - Dimensionality reduction (PCA/PCoA, t-SNE, UMAP)
    - Clustering

- Supervised machine learning
    - Logistic regression
    - Support vector machines
    - Decision trees and random forest
    - Neural networks

- Model evaluation
    - Cross-validation
    - Hyperparameter tuning
    - Model evaluation

Course participants will:

- Recognize when machine learning methods may be beneficial for their research.
- Identify common pitfalls in the application of machine learning methods.
- Gain confidence to provide constructive feedback for applications of machine learning in the manuscripts they review.
- Evaluate the strengths and weaknesses of machine learning approaches presented in the scientific literature.
- Gain familiarity with additional resources to deepen their understanding of machine learning.

Prerequisites: All participants will need to bring a laptop to perform example exercises during the course. Basic knowledge of mass spectrometry and data analysis is recommended. Basic knowledge of Python is recommended. No extensive machine learning experience is needed.