

Machine Learning for the Social Science

Overview:

This free four-day course aims to introduce machine learning techniques and their applications in social science research. Through a combination of lectures and paper discussions, participants will gain a thorough conceptual understanding of implementing machine learning algorithms in social science research.

Course Schedule:

<u>Day</u>	<u>Theme</u>	<u>Lecture</u>	<u>Paper Discussion</u>
1	Health	Introduction to Machine Learning (ML) and Regularized Regression (RR)	Mortality prediction using RR; classification metrics
2	Education	Flexible models: Support Vector Machines (SVM) and ensemble learning	Predicting student graduation using SVM; ensemble methods
3	Crime	Tree-based models and Random Forests (RF)	Predicting recidivism in homicide using Classification Trees; ethical and social implications
4	Health	Artificial Neural Networks (ANN)	Breast cancer classification using ANN; backpropagation and gradient descent

Lecturers:

Ilja Cornelisz (e-mail: i.cornelisz@vu.nl)

Chris van Klaveren (e-mail: c.p.b.j.van.klaveren@vu.nl)

Practical Information:

This is a foundational intermediate course in supervised learning, with a focus on regression and classification methods. We introduce methods intuitively to foster understanding, complemented with the corresponding formulas and derivations, so expect some mathematics in the lectures. As a prerequisite for the course, participants are expected to have a solid understanding of basic regression analysis techniques. If you are taking this course as part of your Training & Supervision Plan (TSP), please note that this course corresponds to 1 ECTS (i.e., 28 hours) of study load.

Each course day is scheduled from 10:00-16:00 hours (allowing for lunch breaks and short coffee breaks). The minimum number of participants for the course is set at 15 and is capped at a maximum of 40 participants. If necessary, a system of first-come first-served applies.