



Benchmarking in the social sciences

Paulina Pankowska
(Utrecht University)

01.02.2023, 16.00 (c.t.)

Department of Statistics, Ludwigstr. 33, Room 144
and online via Zoom (Link)
(Meeting-ID: 913-2473-4411; Password: StatsCol22) a

Benchmarking is about creating a system of reference, in which various models that address the same task can be compared, in order to be able to analyse the strength and weaknesses of various approaches and gain new insights. For computer scientists, benchmarks provide a standardized way to validate predictive models that has proven to be able to detect major breakthroughs, like deep learning.

A benchmarking challenge is when various teams of participants are set a challenge to predict a particular outcome and then their performance is evaluated based on a predefined set of matrices and evaluation criteria. Participants are provided with a training dataset which includes a target variable (dependent variable) and a range of predictor variables (independent variables). They then train a model on this dataset that can then predict the target variable based on values of the predictors. The resulting model is evaluated by predicting values in a holdout dataset. This is a dataset containing typically around 20% of observations which the participants have not had access to. The winner of the challenge is the group which can most effectively predict the target variable in the holdout data.

While benchmarking challenges are popular in data science, machine learning and other fields, there have been very few benchmark challenges in the social sciences. Therefore, during the 2022 ODISSEI-SICSS Summer School, we organized a social science benchmark, which included approx. 20 participants, who were divided into six groups. The aim of the challenge was to predict precarious employment in 2020, based on predictors from 2010 or earlier.

Biography:

Paulina Pankowska is an Assistant Professor at the Sociology Department of Utrecht University. She is currently leading a project that aims to design and set-up a social science benchmark. Her research relates primarily to data and methods quality in the social sciences. She has also published in the field of official statistics, survey research methods, and machine learning.