Artificial Intelligence to create maps

Can we make machine learning algorithms explicitly spatial?

Department:	Physical Geography
Research group:	Computational Geography
Supervisor:	Madlene Nussbaum, Derek Karssenberg
Email address:	m.nussbaum@uu.nl

Project description

In numerous domains, the creation of maps relies on spatial prediction. Observations, such as rainfall levels, air pollutant concentrations, or soil pH, are typically available only at discrete point locations. To generate continuous maps for e. g. arable fields, entire regions, countries, or even the entire globe, one uses other spatial predictor data like satellite images or elevation models and also tries to exploit "similarities" in the data. According to a fundamental principle of geography observations closer to each other are often more similar than observations further away. The extent to which close observations are similar, and the distance over which this relationship holds, is explicitly quantified through the concept of spatial auto-correlation using a kriging variogram.

Machine learning (ML) techniques were not initially designed for spatial mapping and lack inherent knowledge of the relative locations of observations. Nevertheless, ML has become the predominant choice for spatial mapping due to its user-friendliness and good performance, particularly with large datasets featuring numerous predictor variables. Consequently, many studies resort to unsatisfactory ad-hoc methods to incorporate location information. However, ML algorithms could potentially directly account for spatial autocorrelation by integrating the variogram function into the learning process of the ML algorithm. The implementation of such a methodology, referred to here as ML Kriging, is the main focus of this BMA.

We are not aware of any application of "ML Kriging" to create prediction maps. We are therefore highly interested to know:

- How does ML kriging need to be configured for spatial mapping (settings for the variogram function, smoothness of predictions)?
- How does ML kriging perform for a small study area like an arable field where no additional information is available? How well does it perform compared to classical kriging?
- How can ML kriging be applied for larger study areas when also predictors are included?
- Is ML kriging finally the solution to make ML truly spatial?

To investigate these questions datasets from different domains will be analyzed. If a candidate works on spatial mapping within his/her master thesis this dataset could be used. You will execute the research in cooperation with experts in AI, geostatistics, and research software engineering from the Computational Geography research group. Results potentially lead to a peer-reviewed publication.

Job requirements

At least basic knowledge of R (software is not available in Python), either some background in ML or classical geostatistics (kriging methods). General background in statistics and interest to learn more.