Optimization of Training Data Distribution for Regression-based Machine Learning

Methods for Selecting Data Points when Learning or Approximating a Target Mapping

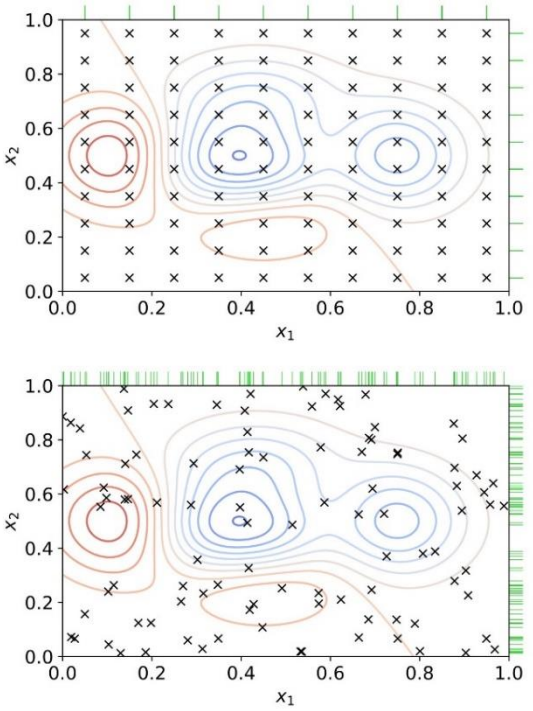
TECHNISCHE UNIVERSITÄT DARMSTADT

Master Thesis

In various use cases in control engineering, it is necessary to learn a possible complex and high-dimensional function based on data samples. Typical applications are the creation of a simple replacement function for a complicated context or the necessity to perform optimizations in the high-dimensional function space (e.g., in hyperparameter optimization in the field of Machine Learning). In the context of control engineering, applications could be the approximation of either the dynamic equations, the controller or the state estimator of the system. To learn the data-based (substitute) function, various methods are available on how to learn the unknown objective function based on the available data.

In this master thesis, the question of how to optimally select the data samples with which we can approximate a static mapping should be explored. For this, a literature research on existing techniques, such as grid, random, and Bayesian search must be conducted. Based on these results the methods should be evaluated regarding data efficiency, resulting approximation quality, and computational cost. The evaluation should be done while considering prominent machine learning algorithms ,e.g., Neural Networks and Gaussian Processes.

If time allows, more complex methods for sampling and approximation of dynamic mappings, such as Event-Based Sampling and Recurrent Neural Networks, can be investigated by the student.



Source: Alexander Elvers - Wikipedia


M.Sc. Hendrik Alsmeier
M.Sc. Felix Häusser
Tel.: 0391 / 67-58776
E-Mail: hendrik.alsmeier@iat.tu-darmstadt.de
felix.haeusser@iat.tu-darmstadt.de
Web: https://www.ccps.tu-darmstadt.de/ccps/

CCPS Control and Cyber-Physical Systems