Contribution ID: **12** Type: **not specified**

Model predictivity assessment: incremental test-set selection and accuracy evaluation

Tuesday, June 28, 2022 2:40 PM (30 minutes)

Unbiased assessment of the predictivity of models learnt by supervised machine-learning methods requires knowledge of the learnt function over a reserved test set (not used by the learning algorithm). Indeed, some industrial context requires the model predictivity to be estimated on a test set strictly disjoint from the learning set, which excludes cross-validation techniques. The quality of the assessment depends, naturally, on the selected test set and on the error statistic used to estimate the prediction error.
In this work we tackle both issues, first by using incremental experimental design methods to "optimally" select the test points on which the criterion is computed. Second, we propose a new predictivity criterion that carefully weights the individual observed errors to obtain a global error estimate. Several incremental constructions are studied. We start with the fully-sequential space filling design (selecting a new point as far away as possible from the existing design). Additionally, we study the support points and kernel herding that are based on the iterative minimization of the Maximum Mean Discrepancy (MMD) between the input distribution and the empirical distribution of different kernels.
Our results show that the incremental and weighted versions of the latter two, based on MMD concepts, yield superior performance. An industrial use case in the domain of nuclear safety assessment, concerning the simulation of thermal-hydraulic phenomena inside nuclear pressurized water reactors illustrates the practical relevance of this methodology, indicating that it is an efficient alternative to cross-validation.

Keywords

Design of experiments, Metamodel, Validation

Primary authors: FEKHARI, Elias (EDF R&D); IOOSS, Bertrand (EDF R&D); MURÉ, Joseph (EDF R&D); PRONZATO, Luc (CNRS, Université Côte d'Azur, Laboratoire I3S); RENDAS, Maria Joao (CNRS, Université Côte d'Azur, Laboratoire I3S)

Presenter: FEKHARI, Elias (EDF R&D)

Session Classification: INVITED SFdS

Track Classification: Other/special session/invited session