



Contribution ID: 8

Type: **not specified**

Experimental Design for Missing Physics

Model-based approaches are commonly used in the analysis, control and optimization of biosystems. These models rely on knowledge of physical, chemical and biological laws, such as conservation laws, transport phenomena and reaction kinetics, which are usually described by a system of non-linear differential equations.

Often our knowledge of the laws acting on the system is incomplete. These gaps in our knowledge are also referred to as missing physics. Experimental data can be used to fill in such missing physics.

Universal Differential Equations (UDE) have recently been proposed to learn the missing parts of the structure. These UDE use neural networks to represent terms of the model for which the underlying structure is unknown.

Because the opaque nature of neural networks is often not desirable in a scientific computing setting, UDE based techniques are often combined with interpretable machine learning techniques, such as symbolic regression. These techniques post-process the neural network to a human-readable model structure.

Because neural networks are data-hungry, it is important that these applications gather highly informative data. However, current model based design of experiment (MbDoE) methodology focuses on parameter precision or discriminating between a finite number of possible model structures. When part of the model structure is entirely unknown, neither of these techniques can be directly applied.

In this presentation, we propose an efficient data gathering technique for filling in missing physics with a universal differential equation, made interpretable with symbolic regression.

More specifically, a sequential experimental design technique is developed, where an experiment is performed to discriminate between the plausible model structures suggested by symbolic regression. The new data is then used to retrain the UDE, which leads to a new set of plausible model structures by applying symbolic regression again.

This methodology is applied to a bioreactor, and is shown to perform better than a randomly controlled experiment, as showcased here:

https://docs.sciml.ai/Overview/dev/showcase/optimal_data_gathering_for_missing_physics/

Special/ Invited session

Classification

Mainly methodology

Keywords

Universal Differential Equation, Symbolic Regression, Model Discrimination

Primary author: STROUWEN, Arno (Strouwen Statistics; PumasAI; KULeuven)

Co-author: Mr MICLUȚA-CÂMPEANU, Sebastian (JuliaHub; University of Bucharest)

Presenter: STROUWEN, Arno (Strouwen Statistics; PumasAI; KULeuven)

Track Classification: Design of Experiments