HAMLET
How to Apply Machine Learning to Experimental & Theoretical PHYSICS
August 19 - 21, 2024
Copenhagen, Denmark

Contribution ID: **35** Type: **Plenary**

Symbolic regression for Science: challenges and opportunities.

Wednesday 21 August 2024 09:30 (25 minutes)

Symbolic regression (SR) is a machine learning technique that seeks to identify mathematical expressions that best fit a given dataset. This method is particularly beneficial in scientific research, due to its ability to generate interpretable models that provide insight into the underlying mechanisms, in sharp contrast with purely data-driven, black-box models. Interpretable machine learning is gaining importance as models are increasingly deployed in critical societal applications like finance and medicine. Also, models discovered by SR are prone to analyses that allow to assess their out-of-distribution behavior. These beneficial properties have led to the application of SR across various fields, including physics, biology, climate modeling, finance, and various engineering disciplines. Despite these opportunities, the application of symbolic regression faces several challenges. One major issue is the computational complexity associated to the combinatorial search within the space of mathematical expressions, which requires the use of high-performance computing, scalable algorithms, and smart strategies or integration of domain-specific knowledge to guide the search and make it more efficient. Additionally, ensuring robustness and generalization of the discovered models can be difficult, as overfitting to noisy data or limited datasets may occur. In this talk, we wish to stimulate the debate among researchers about both opportunities and challenges concerning the use of SR for Science.

Author: LUCANTONIO, Alessandro (Aarhus University)

Presenter: LUCANTONIO, Alessandro (Aarhus University)

Session Classification: Plenaries & Keynotes