
Contribution ID: **97** Type: **Poster**

# Exhaustive Symbolic Regression: Learning Astrophysics directly from Data

A key challenge in the field of AI is to make machine-assisted discovery interpretable, enabling it not only to uncover correlations but also to improve our physical understanding of the world. A nascent branch of machine learning – Symbolic Regression (SR) – aims to discover the optimal functional representations of datasets, producing perfectly interpretable outputs (equations) by construction. SR is traditionally done using a "genetic algorithm"which stochastically selects trial functions by analogy with natural selection; I will describe the more ambitious approach of exhaustively searching and evaluating function space.

Coupled to an information-theoretic model selection principle based on minimum description length, our algorithm "Exhaustive Symbolic Regression" (ESR) is guaranteed to find the simple functions that optimally balance accuracy with simplicity on a dataset. This gives it broad application across science. I will detail the method, its relation to Bayesian statistics and an optional language model-based prior on functions designed to enhance their physicality. Then I will use ESR to quantify the extent to which state-of-the-art astrophysical theories – FLRW cosmology, General Relativity and Inflation – are implied by the current data.

**Presenter:** DESMOND, Harry (University of Portsmouth)

**Session Classification:** Social