Contribution ID: **9**   Type: **not specified**

Structure-Induced Interpretability in Kolmogorov–Arnold Networks

Monday 8 December 2025 09:00 (45 minutes)

Kolmogorov-Arnold Networks exhibit several properties beneficial to scientific discovery. I will outline these properties and show how they can be leveraged to build more interpretable AI models, which I will illustrate with an example from representation theory. I will also discuss how the intrinsic structure of KANs facilitates symbolic regression by employing ideas similar to Google's fun-search. Incorporating LLMs and vision transformers into the regression workflow allows for priming the NN with domain-specific symbolic regression targets. Finally, I will describe how KAN's intrinsic structure can be thought of as arising from wide, dense neural networks that have been sparsified in a specific way, and I will speculate on how this structural perspective relates to desirable empirical properties of KANs.

Presenter: RUHLE, Fabian