INTERPRETABLE FINE-TUNING AND ERROR INDICATION FOR GRAPH NEURAL NETWORK SURROGATE MODELS

Romit Maulik,[1] Hojin Kim,[1] and Shivam Barwey[2]

[1]Pennsylvania State University
[2]Argonne National Laboratory

ABSTRACT

Data-driven surrogate modeling has surged in capability in recent years with the emergence of graph neural networks (GNNs), which can operate directly on mesh-based representations of data. The goal of this work is to introduce an interpretable fine-tuning strategy for GNNs, with application to unstructured mesh-based fluid dynamics modeling. The end result is an enhanced fine-tuned model that isolates regions in physical space, corresponding to sub-graphs, that are intrinsically linked to the forecasting task while retaining the predictive capability of the baseline. These structures, identified by the fine-tuned GNNs, are adaptively produced in the forward pass and serve as explainable links between the baseline model architecture, the optimization goal, and known problem-specific physics. Additionally, through a regularization procedure, the fine-tuned GNNs can also be used to identify, during inference, graph nodes that correspond to a majority of the anticipated forecasting error, adding a novel interpretable error-tagging capability to baseline models. Demonstrations are performed using unstructured flow field data sourced from flow over a backward-facing step at high Reynolds numbers, with geometry extrapolations demonstrated for ramp and wall-mounted cube configurations.

REFERENCES

[1] S. Barwey, H. Kim, and R. Maulik, Interpretable Fine-Tuning and Error Indication for Graph Neural Network Surrogate Models, *Arxiv Preprint*, 2311.07548, 2023.

[2] S. Barwey, V. Shankar, V. Vishwanathan, and R. Maulik, Multiscale graph neural network autoencoders for interpretable scientific machine learning, *Journal of Computational Physics*, 495, 112537, 2023.