NEUP Nuclear Energy University Program

U.S. Department of Energy

Establishing the Predictive Credibility of Data Driven Scientific Machine Learning in Nuclear Applications

PI: Xu Wu – North Carolina State University **Collaborators**: N/A

Program: Distinguished Early Career Program

ABSTRACT:

New innovations in machine learning (ML) are beginning to drive advances in nuclear engineering (NE), but the full potential of these techniques for data-driven scientific machine learning (SciML) has yet to be fully realized. One barrier is that existing ML methods often do not meet the needs of NE applications. Application-agnostic algorithms, or those designed for more traditional, error-tolerant ML applications such as computer vision and natural language processing, cannot typically be directly applied to scientific data in high-consequence NE problems without non-trivial, task-specific modifications. There are significant gaps in the predictive credibility assessment and improvement of ML models that prevent us from building more trustworthy applications in nuclear reactors. To enable wide acceptance of ML by the nuclear regulators, stakeholders, and policy/decision makers, significant scientific advancements to resolve these gaps are imperative with very demanding timelines. Considering that ML trustworthiness is a very broad area of research, in this project we will only focus on UQ of SciML and the data scarcity issues, with the goal of improving SciML accuracy and establish confidence in SciML predictions.

The main objective of this project is to establish and enhance the predictive credibility of data-driven SciML for nuclear applications through rigorous uncertainty quantification (UQ) of SciML models to establish confidence, as well as deep generative learning to address the data scarcity issue. We will develop innovative and efficient UQ methods to quantify the approximation uncertainties in SciML models, especially in generalized domains where the SciML models are extrapolated. We will develop deep generative learning-based SciML models that can approximate the complicated probability distributions underlying the real data and generate synthetic data to significantly augment the dataset to alleviate the data scarcity issue in many nuclear problems. Finally, the project will augment the applications of SciML in NE scientific computing and prepare the students for transformative solutions across various U.S. Department of Energy missions.

These objectives can be achieved through three Research Thrust areas: (1) UQ of SciML models to establish confidence in ML predictions (Monte Carlo dropout, deep ensemble, variational inference and gradient-based Bayesian inference method to train Bayesian neural networks), (2) deep generative learning for data augmentation to alleviate the data scarcity issue (GANs, VAEs, deep rendering model, normalizing flow, transfer learning and diffusion model-based deep generative models), and (3) develop open source tools/libraries for UQ of SciML and deep generative learning techniques. Furthermore, six Education Thrust areas will be pursued for research/education integration: (1) develop new curriculum on SciML – two courses on fundamental and advanced SciML, (2) complete a new textbook on "SciML with NE Applications" with CRC Press, (3) establish the ARTISANS (ARTificial Intelligence for Simulation of Advanced Nuclear Systems) research group that focuses on UQ, SciML and deep generative learning, (4) organize workshops on UQ and SciML in academic conferences and national labs, (5) promote diversity, equity and inclusion in UQ and SciML research (Young Investigators summer program, IAEA Lise Meitner Program, etc.), and (6) offer SciML education to non-traditional students such as nuclear regulators, stakeholders, and policy/decision makers. Finally three Leadership Thrust areas will be pursued to achieve the project goals: (1) develop AI/ML benchmarks for scientific computing in NE with OECD/NEA, (2) promote SciML in academic conferences (AI/ML tracks, special sessions and panels), and (3) promote SciML in professional societies.