Contribution ID: **132**   Type: **Standard 15 min talk**

[Remote] Visualizing Loss Landscapes for Scientific Edge Machine Learning

Tuesday 15 October 2024 17:00 (15 minutes)

Characterizing the loss of a neural network can provide insights into local structure (e.g., smoothness of the so-called loss landscape) and global properties of the underlying model (e.g., generalization performance). Inspired by powerful tools from topological data analysis (TDA) for summarizing high-dimensional data, we are developing tools for characterizing the underlying shape (or topology) of loss landscapes. We are now exploring real-time scientific edge machine learning applications (e.g., high energy physics, microscopy) and using our tools to help design models and understand their robustness (e.g., how to quantify and visualize model diversity, how noise or quantization change the loss landscape). In this talk, I will focus on two of our recent collaborations. First, we evaluate how LogicNets ensembles perform on scientific machine learning tasks like the data compression task at the CERN Large Hadron Collider (LHC) Compact Muon Solenoid (CMS) experiment. By quantifying and visualizing the diversity of LogicNets ensembles, we hope to understand when ensembling can improve performance and how to decide which models to include in an ensemble. Second, we look at new physics constrained neural network architectures designed for the rapid fitting of force microscopy data. We visualize loss landscapes and their topology, observing sharp valleys in the loss landscapes of successfully trained models, likely reflecting the physical constraints. In contrast, we observe flatter but shallower basins in the loss landscapes of lower performing models, suggesting that training may be difficult and can fail to find a physically reasonable solution in some cases. These results highlight potential failure modes similar to those observed for other physically constrained architectures.

Focus areas

Primary authors: GENIESSE, Caleb (Lawrence Berkeley National Laboratory); CHEN, Jiaqing (Arizona State University); XIE, Tiankai (Arizona State University)

Co-authors: WENG, Olivia (UC San Diego); QIN, Shuyu (Lehigh University); AGAR, Joshua (Drexel University); KASTNER, Ryan (UC San Diego); MAHONEY, Michael (UC Berkeley); MACIEJEWSKI, Ross (Arizona State University); TRAN, Nhan (Fermi National Accelerator Laboratory); WEBER, Gunther H. (Lawrence Berkeley National Laboratory)

Presenter: GENIESSE, Caleb (Lawrence Berkeley National Laboratory)

Session Classification: Contributed talks