A prediction interval method for uncertainty quantification of regression models

Pei Zhang¹, Siyan Liu¹, Dan Lu¹, Guannan Zhang², Ramanan Sankaran¹
¹Computational Sciences and Engineering Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830
²Computer Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830

Summary

This work proposes a novel prediction interval (PI) method for uncertainty quantification (UQ) of regression neural networks (NNs). The method requires no distributional assumption, does not introduce extra hyper-parameters, and can effectively identify out-of-distribution (OOD) samples and quantify their uncertainty.

We demonstrate the advantages of our method in a toy regression task with non-Gaussian noise and two real-world scientific applications, in comparison with two state-of-the-art UQ methods—a quality-driven (QD) approach (Pearce et al., 2018) and a deep ensemble (DE) method (Lakshminarayanan et al., 2017).

Our Method

For a standard regression task, $y = f_w(x): \mathbb{R}^d \rightarrow \mathbb{R}$, with a given dataset $\mathcal{D}_{\text{train}} = \{(x_i, y_i)\}_{i=1}^n$:

Goal: Learn the function $f_w(x)$ and the prediction intervals (PIs) to quantify the uncertainty of the prediction.

Key idea: Learn $f_w(x)$, the upper and lower bounds of the PI separately using three independent NNs.

Step 1: Train $f_w(x)$ NN with dataset $\mathcal{D}_{\text{train}} = \{(x_i, y_i)\}_{i=1}^n$ using mean squared error (MSE) loss.

Step 2: Obtain two new datasets from $f_w(x)$ NN $\mathcal{D}_{\text{upper}} = \{(x_i, f_w(x_i) - y_i)\}_{i=1}^n$ and $\mathcal{D}_{\text{lower}} = \{(x_i, f_w(x_i) - y_i)\}_{i=1}^n$.

Step 3: Train two new NNs—$u(x)$ and $v(x)$—to represent the upper and lower uncertainty profiles with the two datasets $\mathcal{D}_{\text{upper}}$ and $\mathcal{D}_{\text{lower}}$, respectively, using MSE loss.

Step 4: Find two coefficients $\alpha$ and $\beta$ such that a target percentage of training samples are covered by the PI $\left[ f_w - \beta v, f_w + \alpha u \right]$.

Step 5: Get the final PI as $\left[ f_w - \beta v, f_w + \alpha u \right]$.

PIs are calculated for surrogate models of ten variables from the Earth System Land Model (ELM) simulations. Two metrics—prediction interval coverage probability (PICP) and mean prediction interval width (MPIW)—are used for comparison. A sound PI method should have a PICP close to the target value with a small MPIW. Our method outperforms QD by providing PICP closer to the 90% target and MPIW on average 60% narrow. PICP and MPIW provided by QD are sensitive to the hyper-parameters. Our method does not need hyper-parameter fine tuning but QD does.

An OOD-aware Autoencoder-based Combustion Model

Pls are calculated for an autoencoder NN for datasets of syngas CO/H₂ combustion with 12 thermo-chemical state variables. ID test set is from a 0-D reactor and OOD test set is from a 3-D direct numerical simulation (DNS) of turbulent flames.

DE with a single run in the first row fails to capture the difference in uncertainties for ID and OOD samples. DE with 10 runs in the second row shows improved but still limited separation of OOD samples from ID and it fails to produce the uncertainty-error correlation. Our method in the third row shows a strong correlation between the uncertainty and the error, and clearly demonstrates that OOD and ID have different uncertainty magnitudes.

Acknowledgement

This material is based upon work supported in part by the U.S. Department of Energy, Office of Science, Offices of Advanced Scientific Computing Research, under the contract ERKJ352, and by the AI Initiative at the Oak Ridge National Laboratory (ORNL). ORNL is operated by UT-Battelle, LLC, for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

References