SAND2021-13926C

Sandia National Laboratories

Machine Learning in Computational Science

Habib N. Najm

Sandia National Laboratories, Livermore CA

AI4ESP Neural Network Session

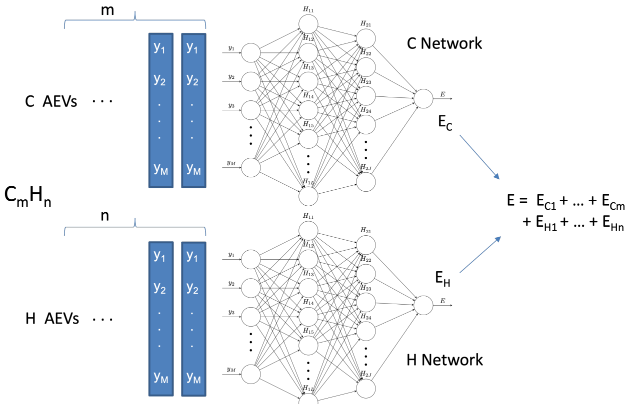
11/02/2021

This work was supported by the US Department of Energy (DOE), Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, & Biosciences (CSGB).

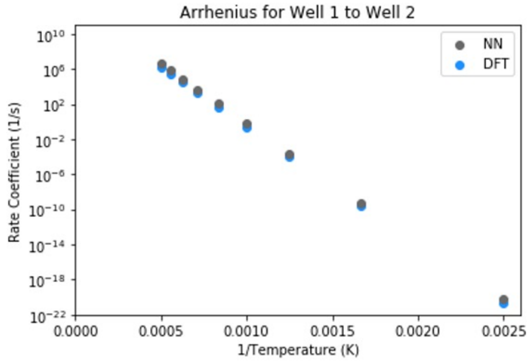
Machine learning has had a broad impact in science

- ML has provided significant advances in predictive science & engineering
  - Flexible accurate DNN surrogates
  - More efficient computations
  - Striking example: protein folding
    - DeepMind's AlphaFold

- E.g. NN potential energy surface (PES)
  - High dimensional complex surface
  - Chemistry, material science
  - PES exploration, MD computations



NN construction for the PES of hydrocarbon molecules



A C_5H_5 system reaction rate computed from NN PES vs DFT

Work with: Yoona Yang, Seamus Hoolahan, Mike Eldred, and Judit Zador, SNL

Utility of Enforcement of Physical Constraints in ML

- Constraints enforcement confines NN predictions to a manifold, which reduces the dimensionality of the optimization problem
  - Expect a simpler loss landscape
  - Reduces NN training effort: exclude exploration in unphysical directions
- Ensures that NN predictions satisfy conservation/symmetries/invariances
- Does not risk error/bias, while enforcement of physical *models* can
  - Physical constraints, conservation laws, invariances are incontrovertible
  - Physical models, relying often on constitutive laws, can be approximate
- Note: role of geometric constraints in superior accuracy of AlphaFold2

Jumper, J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2
https://www.nature.com/articles/d41586-020-03348-4

Means of Enforcement of Physical Constraints in ML

- Explicit enforcement via regularization with governing equation residual
- Implicit enforcement via tailored feature design or NN structure
  - This enforcement *by construction* is in principle preferrable
    - Satisfies constraints irrespective of training
- Examples of implicit enforcement:
  - Feature vector use of *symmetry functions* for NNPES in chemistry Behler & Parrinello 2007
    - Pre-specified uniform-length summary of molecular geometry
  - Alternately, Graph CNNs with internal molecular coordinates Cho & Choi 2019
  - Turbulent stress tensor invariance properties embedded in NN-structure Ling et al 2016
- Challenge:
  - General strategies for enforcement of constraints in governing equations

ML potential in dynamical systems

- Regression
  - Surrogate NN for compute-intensive system subcomponents
  - Data reduction via surrogate NN or reduced-order representations
  - Inputs to (smooth) observables surrogate NN for Bayesian inference
  - Surrogate NN for coarse representation of full model prediction
    - Dynamical features, structural features, …
- Classification
  - Learning from simulation/observational data
  - Discovery of underlying dynamical features
- Learning of optimal features as predictable information-rich summaries of system dynamics