Neural networks learning quantum chemistry

Olexandr Isayev

University of North Carolina at Chapel Hill, Chapel Hill, NC.

Historically, computational chemistry has been unable to overcome the orthogonal requirements of speed and accuracy. If you want an accurate calculation, be ready to pay the price in computer time. In the pursuit of a compromise between speed and accuracy many approximations were developed that expedite *ab-initio* methods with some loss of accuracy. Popular strategies include the partition of the system of interest into fragments, linear scaling, semi-empirical methods or the construction of empirical potentials that have been parametrized to reproduce experimental or accurate *ab-initio* data.

In this talk, we will present a fully transferable deep learning potential that is applicable to complex and diverse molecular systems well beyond the training dataset. Recently we introduced ANAKIN-ME (Accurate NeurAl networK engINe for Molecular Energies) or ANI in short. ANI is a new *method and sampling procedure* for training neural network potentials that utilizes a special kind of symmetry functions to build single-atom atomic environment vectors (AEV) as a molecular representation.

The AI methods that focuses on the use of large and diverse data sets in training new potentials, has consistently proven to be universally applicable to systems containing the atomic species in the training set. Focusing on parametrization for organic molecules (with CHNOSFCl atoms so far), we have developed a universal neural network potential which is *highly accurate compared to reference QM calculations at speeds 10^7 faster*. The potential is shown to accurately represent the underlying physical chemistry of molecules through various test cases including: *chemical reactions (both thermodynamics and kinetics), thermochemistry, structural optimization, and molecular dynamics simulations*. The results presented in this talk will provide evidence of the universal applicability of deep learning to various chemistry problems involving organic molecules.

References:

1. J. S. Smith, O. Isayev, A. E. Roitberg. ANI-1: An extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.*, **2017**, 8, 3192-3203.
2. J. S. Smith, B. Nebgen, N. Lubbers, O. Isayev, A. E. Roitberg. Less is more: sampling chemical space with active learning. *J. Chem. Phys.* **2018**. 148, 241733.
3. K. T Butler, D. W Davies, H. Cartwright, O. Isayev, A. Walsh. Machine learning for molecular and materials science. *Nature*. **2018**, 559, 547–555.
4. J. S. Smith, B. T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev*, A. Roitberg. Outsmarting Quantum Chemistry Through Transfer Learning. *Nature Comm.*, **2019**, Accepted. Preprint: https://chemrxiv.org/articles/Outsmarting_Quantum_Chemistry_Through_Transfer_Learning/6744440