ANAKIN-ME: Using deep learning to develop a fully-transferable and chemically accurate GPU-accelerated potential

Adrian E. Roitberg

Department of Chemistry, University of Florida, Gaiensville, FL, USA
roitberg@ufl.edu

In the field of theoretical chemistry, a compromise between speed and accuracy is required to study the energetics of chemical systems. Quantum mechanical (QM) methods allow accurate energy calculations but at a large computational cost. Classical force fields are very fast, but only accurate near equilibrium and typically incapable of describing chemical reactions. We present the development of a machine learning potential that utilizes deep learning to produce neural network potentials (NNP) from a very large data set of QM reference energies. Through the development of a new methodology, which we call ANAKIN-ME (ANI), this new class of NNP is fully transferable and has chemical accuracy (1 kcal/mol) for organic molecules, containing H, C, N, and O in our work, regardless of the molecular size. To use this ground-breaking methodology, we develop a NVIDIA GPU accelerated software suite, written in C++, Python, and CUDA, called NeuroChem, to train, validate, and test ANI model potentials, as well as predict molecular energies and forces. The use of GPU acceleration allows an ANI potential to be fully trained on a data set containing tens of millions of conformations of molecules with eight or fewer heavy atoms in less than two days. We provide evidence that our method produces potentials that can predict energies and forces of molecules larger than those included in the training set through comparisons to density functional theory (DFT) reference data. We take advantage of the inherently parallel nature of neural networks potentials through GPU acceleration, which allows the user to compute total energies within chemical accuracy of the reference QM data set in roughly 107 less time.

https://roitberg.chem.ufl.edu