ANAKIN-ME: Development of a general purpose neural network potential

Adrian E. Roitberg
University of Florida

In theoretical chemistry, a compromise between speed and accuracy is required to study the energetics of chemical systems, especially with regards to molecular dynamics simulation. Quantum mechanical (QM) methods allow accurate energy calculations with proper physics considerations but can require a massive computational effort. Classical force fields are very fast but only accurate near equilibrium and fail completely in reactivity studies due to a classical physics functional form. One solution to these problems is the development of empirical potentials (EP) built through machine learning (ML) methods. A most notable ML potential method uses artificial neural networks to develop neural network potentials (NNP), which are built by fitting a very flexible, non-physically motivated, continuous function to QM reference energies. Through the development of a new methodology, known as ANAKIN-ME (ANI), we provide the tools to build a new class of NNP that is fully transferable and has chemical accuracy within an entire class of molecules. With the ANI method we develop the ANI-1 general purpose potential for organic molecules. Through extensive benchmark and test cases, the ANI-1 potential provides evidence that the ANI method produces chemically accuracy potentials which predict energies of molecules larger than those included in the training set. Using ANI-1, we compare against a test set containing up to 10 atoms (C, N, O) and show that we can predict total energies with 1.92 kcal/mol RMSE vs wB97X/6-31G*, at a cost roughly six orders of magnitude faster than the full QM calculations. The ANI method brings a new, highly efficient and very accurate, method for the development of NNPs into the realm of reality, and opens the door for a new generation of “out-of-the-box” general purpose potentials.