Accepted Manuscript

Amp: A modular approach to machine learning in atomistic simulations

Alireza Khorshidi, Andrew Peterson

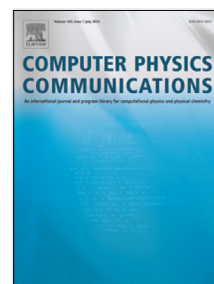Please cite this article as: A. Khorshidi, A. Peterson, *Amp*: A modular approach to machine learning in atomistic simulations, *Computer Physics Communications* (2016), http://dx.doi.org/10.1016/j.cpc.2016.05.010

*Manuscript

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Amp: A modular approach to machine learning in atomistic simulations

Alireza Khorshidi, Andrew Peterson[*]

School of Engineering, Brown University, Providence, Rhode Island, 02912, United States.
*andrew_peterson@brown.edu

Abstract

Electronic structure calculations, such as those employing Kohn-Sham density functional theory or *ab initio* wavefunction theories, have allowed for atomistic-level understandings of a wide variety of phenomena and properties of matter at small scales. However, the computational cost of electronic structure methods drastically increases with length and time scales, which makes these methods difficult for long time-scale molecular dynamics simulations or large-sized systems. Machine-learning techniques can provide accurate potentials that can match the quality of electronic structure calculations, provided sufficient training data. These potentials can then be used to rapidly simulate large and long time-scale phenomena at similar quality to the parent electronic structure approach. Machine-learning potentials usually take a bias-free mathematical form and can be readily developed for a wide variety of systems. Electronic structure calculations have favorable properties—namely that they are noiseless and targeted training data can be produced on-demand—that make them particularly well-suited for machine learning. This paper discusses our modular approach to atomistic machine learning through the development of the open-source Atomistic Machine-learning Package (*Amp*), which allows for representations of both the total and atom-centered potential energy surface, in both periodic and non-periodic systems. Potentials developed through the atom-centered approach are simultaneously applicable for systems with various sizes. Interpolation can be enhanced by introducing custom descriptors of the local environment. We demonstrate this in the current work for Gaussian-type, bispectrum, and Zernike-type descriptors. *Amp* has an intuitive and modular structure with an interface through the python scripting language yet has parallelizable fortran components for demanding tasks; it is designed to integrate closely with the widely used Atomic Simulation Environment (ASE), which makes it compatible with a wide variety of commercial and open-source electronic structure codes. We finally demonstrate that the neural network model inside *Amp* can accurately interpolate electronic structure energies as well as forces of thousands of multi-species atomic systems.

Program Summary

Program title: *Amp*