**INTRODUCTION**

- **Significance**: Representing the inter-atomic potential energy surface (PES), both accurately and efficiently, is one of the most important and challenging problems in molecular modeling.
- **Traditional dilemma**: *Ab initio* methods are accurate but highly inefficient, while empirical force fields are efficient, but there is a limited guarantee for their accuracy.
- **Deep learning may help**: Neural networks are capable of learning complex, non-linear, and many-body functional dependence.
- **PES as regression**: There have been many interesting and challenging large-scale datasets for classification tasks, but relatively few for regression. The PES provides a natural candidate for a challenging regression task.

**Goal**

We wish to have a potential energy model that is

- **Accurate** for both finite and extended systems, i.e., able to reproduce quantum mechanics results with high fidelity.
- **Efficient**, several orders of magnitude faster than quantum mechanics, hopefully with linear scaling.
- **End-to-end**, human intervention should be minimized. The only input should be the chemical species and the atomic coordinates.
- **Symmetry preserving**, invariant under translational, rotational, and permutational operations.
- **Size extensive and smooth**.

**CONTRIBUTION**

1. We propose Deep Potential–Smooth Edition (DeepPot-SE) model that satisfies all the requirements listed above. The methodology is also applicable to other machine learning tasks that have symmetry-preserving requirements.
2. We test the DeepPot-SE model on various systems, which extend previous studies by incorporating DFT data for challenging materials such as high entropy alloys (HEAs).

**Mathematical Formulation**

**General decomposition**: Given the coordinate matrix \( \mathcal{R} \in \mathbb{R}^{N \times 3} \) of \( N \) atoms: \( \mathcal{R} = [\mathbf{r}_1, \cdots, \mathbf{r}_N] \), define the local environment of atom \( i \) as \( \mathcal{R}' = [\mathbf{r}'_{i1}, \cdots, \mathbf{r}'_{ij}, \cdots, \mathbf{r}'_{iN}] \), and represent the total potential energy as

\[
E^w(\mathcal{R}) = \sum_i E^w(\mathcal{R}_i) = \sum_i E_i.
\]

The DeepPot-SE sub-networks:

1. Map the relative coordinates \( \mathcal{R}' \in \mathbb{R}^{N \times 3} \) onto generalized coordinates \( \mathcal{R}'' \in \mathbb{R}^{N \times 3} \), as \( \{x_{ij}, y_{ij}, z_{ij}\} \rightarrow \{s(r_{ij}), \tilde{x}_{ij}, \tilde{y}_{ij}, \tilde{z}_{ij}\} \),

2. Define the local embedding matrix \( \mathcal{G}' \in \mathbb{R}^{N \times M} \) as

\[
(\mathcal{G}')_d = (G(s(r_{ij})))_d,
\]

where \( G \) is the local embedding depending on the chemical species of both atom \( i \) and its neighbor atom \( j \).

3. Define the encoded feature matrix \( \mathcal{D}' \in \mathbb{R}^{N \times M} \) of atom \( i \)

\[
\mathcal{D}' = (\mathcal{G}')^T \mathcal{R}'' \mathcal{R}'^T (\mathcal{G}')^2
\]

that preserves all the symmetries.

4. Maps the encoded feature matrix \( \mathcal{D}' \) onto the local energy \( E_i \) through another feedforward network.

**FURTHER INFORMATION**

- **Code** [github.com/deepmodeling/](https://github.com/deepmodeling/)
- **Data** [www.deepmd.org/database/deeppot-se-data/](http://www.deepmd.org/database/deeppot-se-data/)
- **Contact** jiequn@princeton.edu, linfengz@gmail.com, han.wang.caep@gmail.com, alsaidi@pitt.edu