

二元合金系に対する高次元ニューラルネットワークポテンシャルの構築
Construction of High-dimensional Neural Network Potentials for Binary Alloys

物理学専攻 修士2年 猿谷隆祐
Ryusuke Saruya

1. 研究の目的

密度汎関数理論に基づいた第一原理計算や第一原理に基づいた分子動力学法により物質の構造や電子状態について解析することが出来る。しかし、計算コストが非常に高いことが大きな問題であり、様々な物質について網羅的に研究することは困難である。この問題に対して機械学習を利用したポテンシャルエネルギーの計算が注目されている。特にニューラルネットワークポテンシャルは、第一原理計算と同程度の高い精度を満たしながら計算コストを大幅に抑えられる点で注目されており、様々な物質について研究が報告されている。[6-9]

本研究では Behler と Parrinello により提案された手法[1-4]を用いて Au-Li の二元系に対して高次元ニューラルネットワークポテンシャル(HDNNP)を構築する。ネットワークの構築には公開されている構造データ[5]を使用した。また、 $\text{Al}_{13}\text{Fe}_4$ の単位胞内に Pt 原子を一つ置換し構造緩和計算を行った結果[10]を用いて Al-Fe-Pt の三元系に対して HDNNP を構築する。

2. 計算理論

ニューラルネットワークは生物の神経網を模して造られたネットワークであり、入力層、中間層、出力層の3つのレイヤーから構成されている。各レイヤーにはノードと呼ばれる人工的なニューロンが含まれており、レイヤーの数と各レイヤー内のノード数によってニューラルネットワークの構造を定義することが出来る。図1に示すニューラルネットワークの出力値は重み w とバイアス b を用いて次式のように表される。

$$y = f_1^3 \left[\sum_{l=1}^3 w_{l1}^{23} f_l^2 \left\{ \sum_{k=1}^3 w_{kl}^{12} f_k^1 \left(\sum_{j=1}^2 w_{jk}^{01} x_j + b_k^1 \right) + b_l^2 \right\} + b_1^3 \right] \quad (1)$$

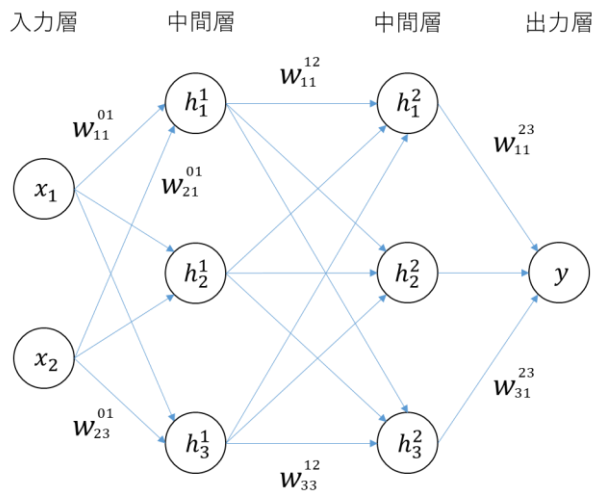


図1: 典型的なニューラルネットワークの構造

図1のような典型的なニューラルネットワークを用いてポテンシャルエネルギーを表現するには大きく3つの問題がある。まず、予測する系の大きさが大きくなるとネットワークが大きくなることでパラメータ数が膨大になり、ネットワークの学習が困難になることである。次に、原子数が異なる系に対応できないという点である。ニューラルネットワークは学習後に入力層内のノード数を変化させることが出来ない。そのため、異なる大きさの単位胞を持つ構造や原子の欠陥が生じた構造など原子数が異なる系に利用することが出来ないという問題が生じる。最後に、ポテンシャルエネルギーの対称性が失われてしまうことである。多くの原子を含む系において同種の原子の位置が入れ替わってもポテンシャルエネルギーは不変である。しかし、ニューラルネットワークは入力層の各ノードごとに重みやバイアスの値が決まっているため、入力されるノードが変化すると出力値も変化する。そのため原子位置の交換に対応して入力ノードの位置を変化させると出力値が変化するため、ポテンシャルエネルギーの対称性を保つことが出来ない。また、結晶の持つ並進対称性や回転対称性を保つことも難しい。

このような問題を解決した手法が Behler と Parrinello により提案された高次元ニューラルネットワーク (HDNNP) である。HDNNP は、原子位置を入力値として系のエネルギーを予測するモデルであり、原子ごとにニューラルネットワーク (Atomic NN) を構築し、各ネットワークの出力値 (Atomic energy) の合計で系のエネルギーを表すモデルである。HDNNP の大きな特徴は、元素ごとに同じ Atomic NN を用意することである。これにより同種の原子の交換に対する普遍性を保つことが出来る。さらに、原子数に合わせて Atomic NN の数を調整することが出来るため、系の原子数の増減に対応できるとともに系の大きさに関わらずパラメータ数を抑えることが出来る。

HDNNP の入力値として原子の周辺環境を記述する対称性関数(symmetry function)の値を用いる。2種類の対称性関数が定義され、(2)式で定義される関数を radial symmetry function と呼び、(3)式で定義される関数を angular symmetry function と呼ぶ。

$$G_i^2 = \sum_{j=1}^N e^{-\eta R_{ij}^2} f_c(R_{ij}) \quad (2)$$

$$G_i^3 = 2^{1-\zeta} \sum_{j \neq i} \sum_{k \neq i, j} (1 + \lambda \cos \theta_{ijk})^\zeta \left[e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_c(R_{ij}) f_c(R_{ik}) f_c(R_{jk}) \right] \quad (3)$$

ここで f_c はカットオフ関数であり、(4)式で定義される関数である。 R_{ij} は中心原子 i と近傍原子 j の原子間距離であり、 R_c はカットオフ半径である。(2)式は中心原子とカットオフ半径内の原子との原子間距離に依存する対称性関数であり η はハイパーパラメータである。(3)式は原子間距離と結合角度から計算される対称性関数であり η, ζ はハイパーパラメータである。また、 $\lambda = \pm 1$ を用いて \cos 関数のピークを反転させて様々な結合角度について表現力を保つ。

$$f_c = \begin{cases} 0.5 \left[\cos \left(\frac{\pi R_{ij}}{R_c} \right) + 1 \right] & (R_{ij} \leq R_c) \\ 0 & (R_{ij} > R_c) \end{cases} \quad (4)$$

本研究では python パッケージの Tensorflow[11]及び Keras[12]を用いて HDNNP を構築した。Au-Li の二元系に対する HDNNP の構築には 8696 個の構造データ[5]を使用し、Al-Fe-Pt の三元系に対する HDNNP の構築には 604 個の構造データ[10]を使用した。それぞれ全データの約 90%のデータを訓練データとして使用した。エネルギーの予測精度は平均二乗誤差 RMSE で評価した。

3. 結果

Au-Li の二元系に対する HDNNP の予測値と正解値の比較を図 2 に示す。図 2(a)は訓練データ、(b)はテストデータに対する予測であり予測精度はそれぞれ 5.734 meV/atom, 5.857 meV/atom であった。訓練データとテストデータそれぞれに対して精度よくエネルギー予測が出来ている。散布図内の直線上に予測結果が乗ることからも予測精度の高いネットワークの構築が出来ていることが分かる。

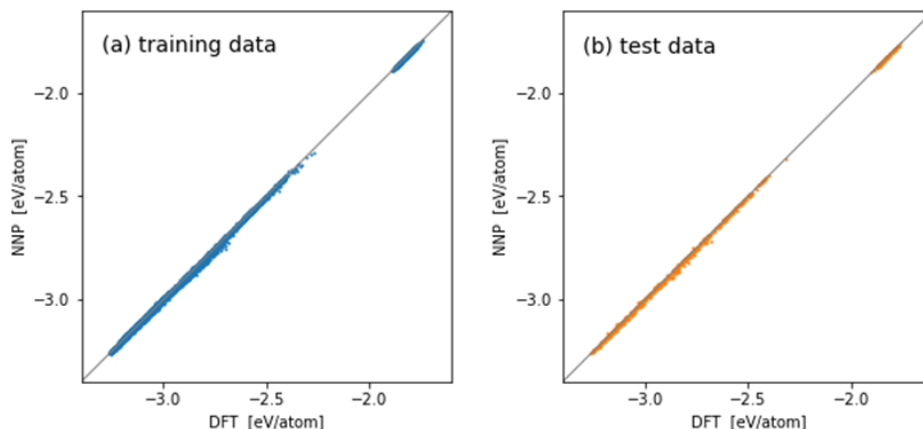


図 2 : DFT 計算と NNP のエネルギー予測

構築したネットワークの汎化能力を評価する。検証データとして 589 個の構造データ [5] を利用し各構造のエネルギーを予測する。予測結果を図 3 に示す。589 個の構造データに対する予測精度は 4.660 meV/atom であった。ニューラルネットワークの構築に全く使用していない構造データに対しても非常に精度よくエネルギー予測が出来ていることから十分な汎化能力があるといえる。

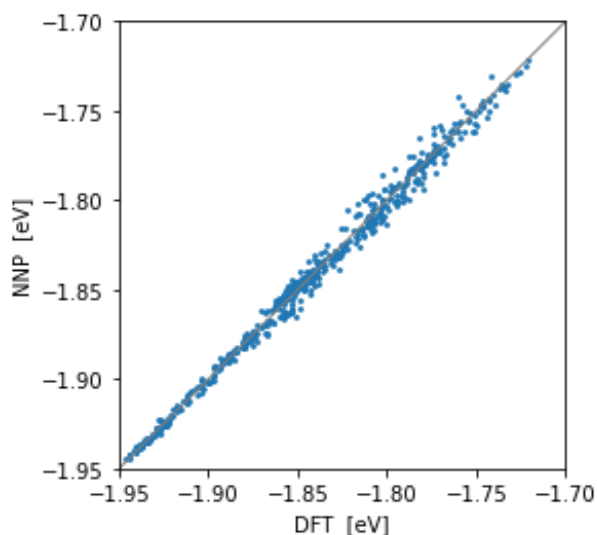


図 3 : 汎化能力の評価

次に、Al-Fe-Pt の三元系に対する HDNNP の予測値と正解値の比較を図 4 に示す。訓練データとテストデータに対する予測精度はそれぞれ 0.416 meV/atom, 0.462 meV/atom であり、予測精度の高いネッ

トワークが構築できた。学習に使用したデータ数が少ないことや似た構造データが多いことから Au-Li の二元系に対するネットワークのような汎化能力は得られておらず、汎化能力の高いネットワークの構築には異なるデータをさらに学習する必要がある。

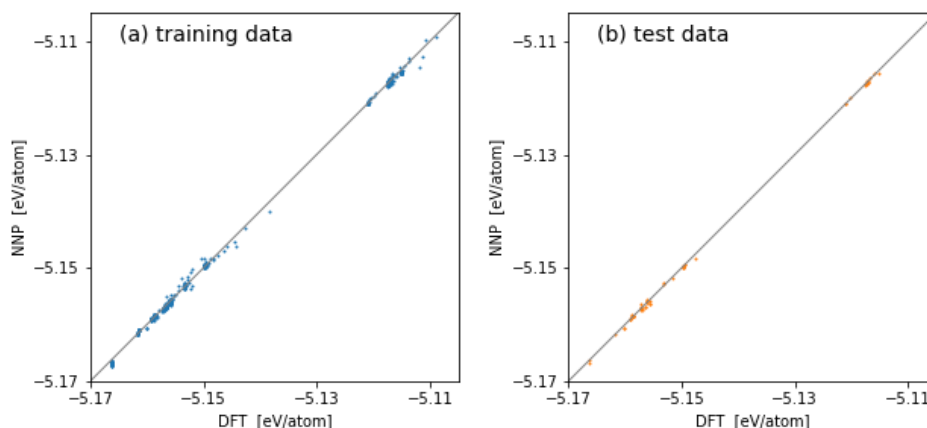


図 4 : DFT 計算と NNP のエネルギー予測

4. 結論

本研究では Au-Li の二元系及び Al-Fe-Pt の三元系に対して高次元ニューラルネットワークポテンシャルを構築し、系のエネルギーを予測するネットワークモデルを構築した。ともに DFT 計算の計算結果を精度よく再現するネットワークであり、未知のデータに対しても汎化能力を持つネットワークの構築が出来た。高次元ニューラルネットワークポテンシャルを利用することで物性研究の効率化に効果を発揮することが期待できる。

汎化能力の高いニューラルネットワークポテンシャルの構築には、適切な構造データセットを学習に用いることが重要であることが示唆される。ネットワークポテンシャルの構築に対する適切な構造データセットの選択や対称性関数のハイパーパラメータの選択に対して確立された手法や理論はないため、こうした問題を解決する手法の開発が求められる。

参考文献

- [1] Jörg Behler and Michele Parrinello, Phys. Rev. Lett. 98, 146401, (2007).
- [2] Jörg Behler, THE JOURNAL OF CHEMICAL PHYSICS 134, 074106, (2011).
- [3] Jörg Behler, International Journal of Quantum Chemistry 2015, 115, 1032–1050.
- [4] Jörg Behler, Chem. Rev. 2021, 121, 10037–10072.
- [5] Koji Shimizu, Elvis F. Arguelles, Wenwen Li, Yasunobu Ando, Emi Minamitani and Satoshi Watanabe, Phys. Rev. B 103, 094112, (2021).
- [6] Nongnuch Artrith, Tobias Morawietz, and Jorg Behler, Phys. Rev. B 83, 153101, (2011).
- [7] Nongnuch Artrith and Jorg Behler, Phys. Rev. B 85, 045439, (2012).
- [8] Marcos del Cueto et al., J. Phys. Chem. C 2020, 124, 5174–5181.
- [9] Abhinav C. P. Jain et al., Phys. Rev. Materials. 5, 053805 (2021).
- [10] 岩戸翔太 鹿児島大学 2023 年度修士論文
- [11] <https://www.tensorflow.org/?hl=ja>
- [12] <https://keras.io>