**Получение аналитической функции потенциала межатомного взаимодействия методами машинного обучения**

*Свирин Денис Витальевич¹, магистр,*

*v.s.denis@yandex.ru*

*Бажанов Дмитрий Игоревич1,2, к.ф.-м.н., доцент,*

[*dima@kintech.ru*](mailto:dima@kintech.ru)

*1МГУ имени М.В. Ломоносова, г. Москва*

*2ФИЦ ИУ РАН, г. Москва*

Аннотация: В данной работе исследуется возможность применения символьной регрессии для получения функции межатомного потенциала взаимодействия в аналитическом виде.

Ключевые слова: межатомный потенциал, машинное обучение, символьная регрессия, теория функционала плотности.

**Введение**

Наиболее точное представление о потенциальной энергии получается из квантово-механических вычислений, однако точное решение уравнения Шредингера для реальных систем недостижимо, поэтому необходимо использовать численные методы для поиска приближенного решения. На сегодняшний день наиболее популярным подходом является теория функционала плотности, которая сочетает скорость и эффективность. Однако этот подход плохо масштабируется и неприменим для больших систем. С ростом интереса к машинному обучению, оно стало применяться и в этой области. В настоящее время существует множество решений использующих алгоритмы машинного обучения, но имеющих различные идеи: в MTP (Moment Tensor Potentials) считается, что можно представить потенциал как линейную комбинацию базисных полиномиальных функций [1]; message-passing networks представляют собой граф имитирующий исследуемую систему [2]; высокоразмерные нейронные потенциалы (High-Dimensional Neural Network Potentials) - комбинация нейросетей аппроксимирующих атомное взаимодействие [3]; Gaussian Approximation Potentials (GAP) используют регрессию гауссовского процесса для моделирования потенциала [4]. Однако и такие методы не лишены недостатков, например, проблемы с интерпретируемостью и необходимостью в большой тренировочной выборке.

Один из интересных подходов машинного обучения позволяющий получить интерпретируемый результат - символьная регрессия [5], который также можно применить и для получения межатомного потенциала.

**Символьная регрессия**

Данный алгоритм позволяет получить аналитическую формулу, описывающую искомые данные (тренировочные). Его работа основывается на генетическом программировании. Данный алгоритм работает итерационно. На каждой итерации создается набор функций, которые представляются в виде деревьев (Рис 1.). Далее идет преобразование данных функций (мутация и скрещивание (Рис. 1)) и получение новых, после этого отбираются N функций, которые наиболее точно описывают тренировочные данные.

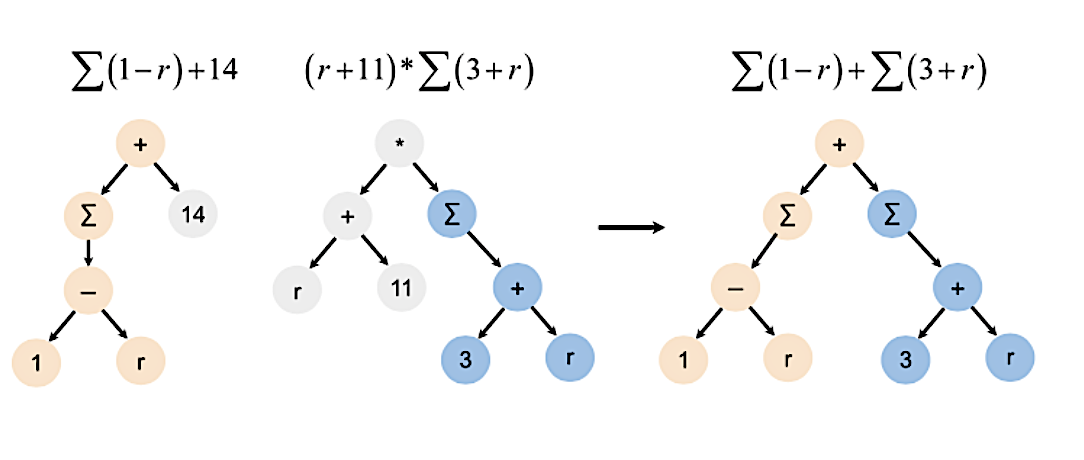


Рис. 1. Пример функций и их возможного преобразования (скрещивание).

В данной работе регрессия будет искать функции вида:

,

где - полная энергия атома ; - функции энергии притяжения и отталкивания для потенциала RGL[6]; - дополнительные функции, - расстояние между атомами и .

Данная форма была основана на полуэмпирическом потенциале RGL [6]. Так как RGL сам по себе способен описывать различные кристаллические системы на основе 3d-5d металлов , то его модификация с помощью функций и поможет сократить сложность модели и задаст вектор развития функций. Более того, использование модификации позволит ускорить время обучения. В роли параметра также выступает и степень . Отклонение от традиционного , которое было получено с помощью теории сильной связи, связано экспериментально обнаруженным плато на которое выходит модель при данном . Задание как свободного параметра не несет серьезной вычислительной нагрузки, так как подобрать оптимальную константу гораздо проще, чем целую функцию.

Так как RGL адаптирован для описания объемных свойств системы, он не показывает хорошего качества на низкоразмерных системах (адатом, димер на поверхности, …)[7]. Поэтому в качестве тренировочной базы были выбраны следующие системы: гранецентрированная кубическая решётка (ГЦК, FCC) *Cu* и адатом *Cu* на поверхности *Cu (001)*. Для формирования тренировочной базы, с помощью теории функционала плотности и программы VASP, были получены группы значений (Рис. 2) для данных двух структур.

**Выводы**

В работе удалось получить аналитическую форму межатомного потенциала способного описывать как объемную, так и низкоразмерную систему:

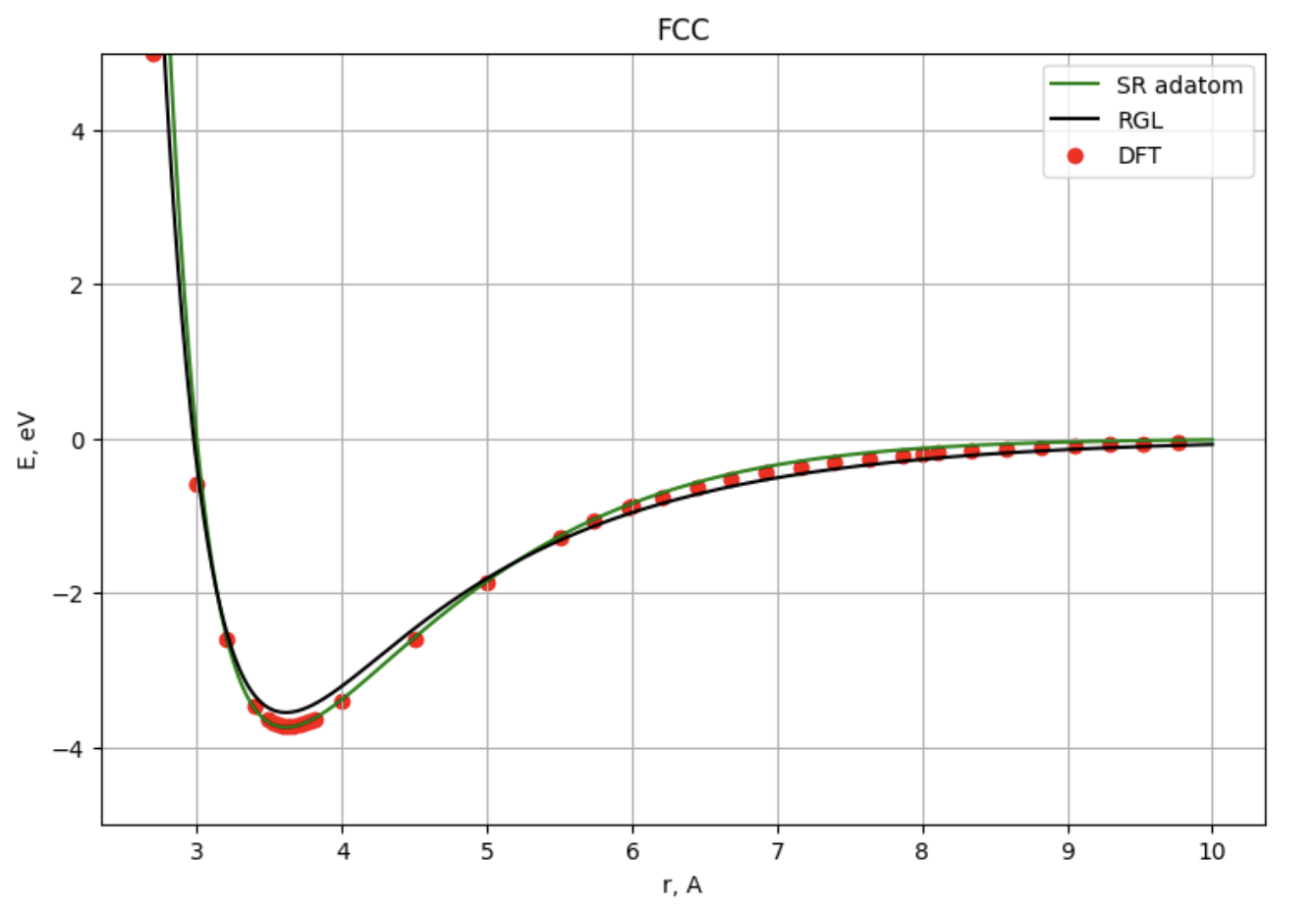
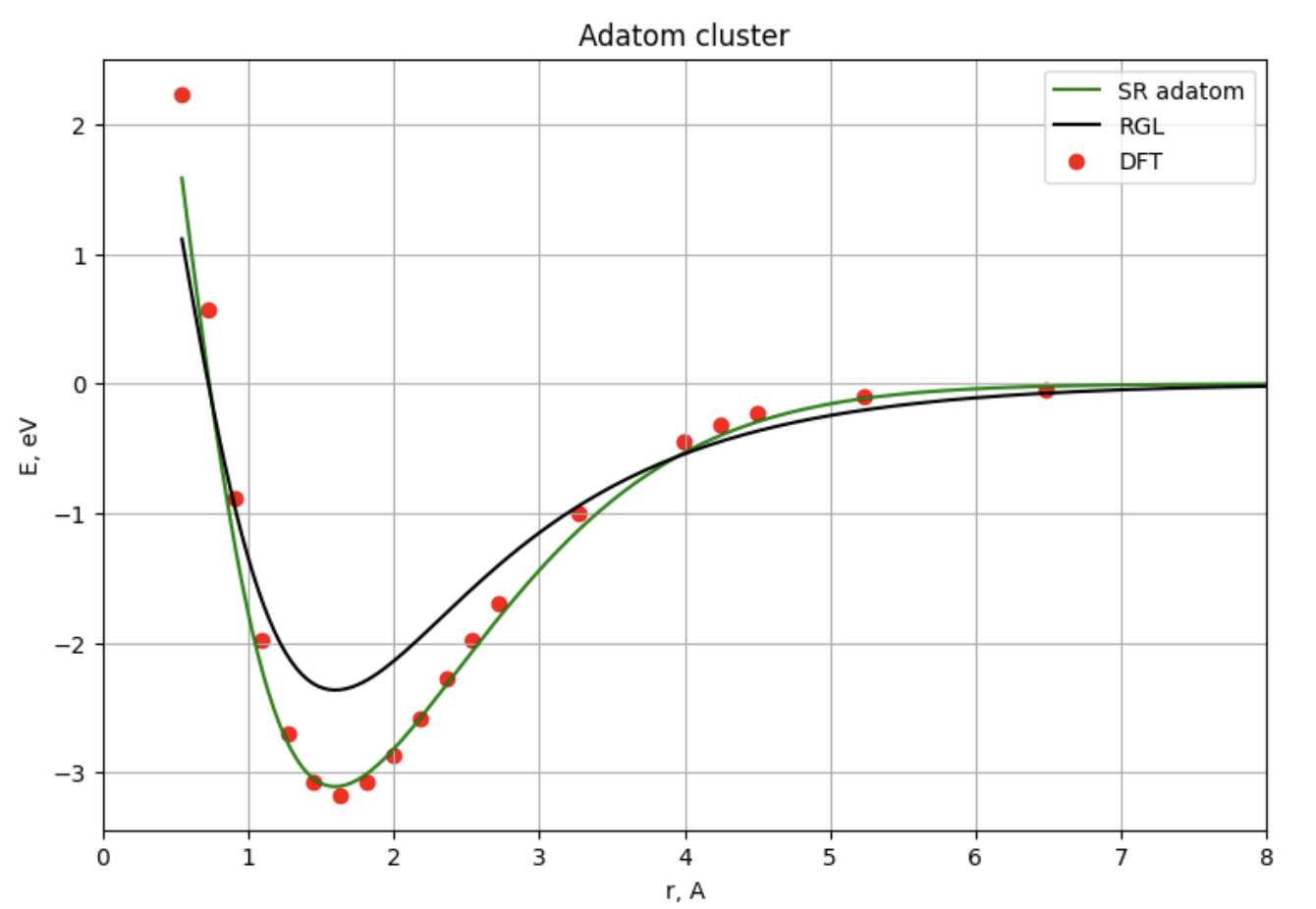
.

Рис.2 Потенциал RGL и результат его модификации.

Данная модификация помогла не только существенно улучшить качество модели при работе с адатомом, но и повысить точность для ГЦК структуры (Рис. 2). Внедрение в роли свободного параметра помогло преодолеть плато, на которое выходила символьная регрессия при . Успешный подгон модели под данные для различных систем показывает потенциальные возможности применения символьной регрессии для получения широко обобщаемого потенциала.

**Список использованных источников**

1. A. Shapeev. Moment tensor potentials: A class of systematically improvable interatomic potentials // Multiscale Model. Simul. 14(3), 1153–1173 (2016)
2. K. T. Schütt, et al. SchNet—A deep learning architecture for molecules and materials //J. Chem. Phys. 148(24), 241722 (2018)
3. J. Behler et al. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces // Phys. Rev. Lett. 98, 146401 (2007)
4. Albert P. Bartók et al. Gaussian approximation potentials: A brief tutorial introduction // Machine Learning and Quantum Mechanics, 2015, Vol. 15, no. 16, pp. 1051-1057
5. T. Mueller et al. Fast, accurate, and transferable many-body interatomic potentials by symbolic regression // [npj Computational Materials](https://www.nature.com/npjcompumats), 5,2019
6. V. Rosato et el. Tight-binding potentials for transition metals and alloys // PHYSICAL REVIEW B, v 48, 1, 1993
7. N. A. Levanov et al. Energetics of Co adatoms on the Cu(001) surface // PHYSICAL REVIEW B, 2000, v61, 3, pp. 2230-2234

**Generating analytical function for interatomic potential using machine learning approac**h

D.V.Svirin, D.I. Bazhanov

Abstract: In this work, symbolic regression is used to generate interatomic potential in analytical form.

Key words: Machine Learning, Symbolic Regression, DFT, Interatomic potential.