

МАШИННОЕ ОБУЧЕНИЕ В КВАНТОВОЙ ХИМИИ

*Измоденов Д.В.⁽¹⁾, Лейбин И.В.⁽¹⁾, Озеров Г.К.⁽²⁾,
Безруков Д.С.^(1,2), Синуцкий А.В.⁽³⁾*

⁽¹⁾ Московский государственный университет
119991, г. Москва, Ленинские горы, д. 1

⁽²⁾ Сколковский институт науки и технологии
121205, г. Москва, Большой бул., д. 30, стр. 1

⁽³⁾ Университет Стэнфорд
94305, США, Калифорния, Стэнфорд

Основной задачей квантовой химии является приближенное решение электронного уравнения Шредингера. Точное аналитическое решение невозможно уже для двухэлектронной системы и на сегодняшний день предложено достаточно большое количество методов приближенного решения уравнения Шредингера. «Золотым стандартом» при расчете энергии основного состояния, за исключением сильно коррелированных систем, считается метод связанных кластеров, однако он является достаточно требовательным к вычислительным мощностям.

Последние несколько лет наблюдается экспоненциальный рост числа работ, посвященный применению техник машинного обучения для приближенного решения уравнения Шредингера. Основным достоинством данного метода является его высокая производительность, обученная нейросеть может предсказывать свойства больших молекулярных систем за времена, сравнимые со временем выполнения молекулярно-механического расчета.

Целью настоящей работы являлось исследование возможностей нейросети архитектуры U-Net, недавно созданной в университете Стэнфорда. Данная нейронная сеть была обучена на базе только органических молекул по результатам расчета функционала плотности B3LYP и дообучена на нескольких сотнях расчетов CCSD/CCSD(T).

На вход сети подается только электронная плотность молекулы, рассчитанная методом Хартри-Фока, а на выходе получается поправка к энергии атомизации и электронной плотности. Задачей настоящей работы являлись тестирование применимости данной нейросети для предсказания свойств неорганических молекул.

В итоге использованная в данной работе нейросеть позволяет предсказывать значения энергий и электронных плотностей даже на неорганических молекулах, которые отсутствовали в обучающей выборке, в том числе значения на молекулах, содержащих элементы, которые не входили в состав молекул из обучающей серии.