

## Нейросети предскажут пользу молекул в 2 тысячи раз быстрее

Ученые из группы «Глубокое обучение в науках о жизни» Института искусственного интеллекта AIRI разработали метод, который позволит симулировать поведение органических молекул в 2 тысячи раз быстрее, чем традиционные подходы, основанные на решении уравнений квантовой физики. Подобные инструменты, исходя из межатомных взаимодействий молекул, дают возможность прогнозировать полезные свойства молекулярных структур без применения физических симуляторов, которые требуют значительных вычислительных мощностей.

Компьютерное моделирование стало незаменимым помощником для химиков, поскольку оно позволяет предсказать свойства молекулы без предварительного синтеза. К числу наиболее популярных численных методов относятся методы на основе теории функционала плотности (DFT), которые позволяют предсказывать энергии молекулярных конформаций с высокой точностью. Однако DFT-симуляторы требуют значительного времени на вычисление.

В последние годы активно развиваются подходы, использующие нейронные сети для более быстрого предсказания молекулярных свойств. Один из них опирается на использование нейросетевых потенциалов (NNP) для предсказания энергии молекулярной конформации. Команда исследователей из AIRI, ФИЦ ИУ РАН, МФТИ и Университета Констрактор в Бремене показала, что оптимизация с использованием NNP примерно в 2 тысячи раз быстрее, чем оптимизация с помощью DFT-симулятора.

Вместе с тем, учёные выяснили, что обученные на имеющихся в открытом доступе наборах данных нейронные потенциалы не могут быть использованы для задачи оптимизации без дообучения. Чтобы получить качество, сравнимое с физическими симуляторами, необходимо собрать и посчитать энергию для примерно полумиллиона дополнительных конформаций.

С целью уменьшить количество необходимых дополнительных данных при обучении нейронного потенциала, исследователи предложили новый фреймворк, названный постепенным выучиванием оптимизации (Gradual Optimization Learning Framework, GOLF). В его основе лежит активное обучение, в котором помимо DFT-симулятора используется суррогатный симулятор на базе простой эмпирической модели молекулярных силовых полей. Эксперименты показали, что нейронный потенциал, обученный с помощью GOLF, имеет такую же точность при в 50 раз меньшем числе дополнительных конформаций. Вместе с тем, научная группа активно занимается развитием других инструментов, полезных для фармацевтической отрасли.

«Традиционно в фармацевтике разработка лекарственного препарата от идеи до выхода на рынок может занимать до 10-20 лет. Нейросети меняют эту ситуацию. Уже сейчас несколько молекул, предложенных нейросетями, проходят клинические испытания, хотя самой идее нейронных потенциалов нет и 10 лет. Мы надеемся, что наши разработки помогут ускорить этот процесс ещё сильнее» — Артур Кадури́н, руководитель группы «Глубокое обучение в науках о жизни».

---

Вопросы: [pr@airi.net](mailto:pr@airi.net)

***Научно-исследовательский Институт искусственного интеллекта AIRI** — автономная некоммерческая организация, занимающаяся фундаментальными и прикладными исследованиями в области искусственного интеллекта. На сегодняшний день более 150 научных сотрудников AIRI задействовано в исследовательских проектах Института для работы совместно с глобальным сообществом разработчиков, академическими и индустриальными партнерами.*