



## Ученые Института Искусственного Интеллекта AIRI создали нейросеть для предсказания эффекта генетических мутаций

Научная группа «Биоинформатика» Института Искусственного Интеллекта AIRI представила новую архитектуру нейронной сети DeepCT, которая поможет врачам устанавливать причины наследственных заболеваний.

Генетики могут установить причину для разных наследственных заболеваний в среднем в 30-40% случаев. Механизмы влияния мутаций на белок-кодирующие регионы ДНК хорошо изучены, однако, такие регионы составляют лишь около 1% всей ДНК человека. Анализ изменений в некодирующих белки (регуляторных) регионах генома с помощью машинного обучения может повысить процент обнаружения причин наследственных заболеваний. В то же время, существующие ML-модели позволяют определить изменение активности гена при появлении изменений в ДНК не во всех типах клеток, а лишь в тех, где необходимые функции были измерены в лаборатории. Таких экспериментальных данных мало, поэтому использовать весь спектр возможностей ИИ для предсказания эффекта мутаций на все ткани организма человека достаточно сложно.

Созданная в AIRI модель обучалась на данных 3 тысяч полногеномных экспериментов для более чем 700 типов клеток человека. В результате модель определила, как разные свойства клеток зависят друг от друга и как на эти свойства влияют изменения в последовательности ДНК. Благодаря этому, если важные для врача характеристики не были измерены в каком-то типе клеток, но для этого типа клеток доступны другие данные, модель сможет предсказать недостающую информацию. В итоге, получив на вход список мутаций, модель может предсказать их последствия для каждого из типов клеток человека.

Для проверки работы нейросети учёные сравнили геномы детей с аутизмом и их здоровых родственников, основываясь на предсказаниях нейросети. Оказалось, что у детей с аутизмом гораздо больше геномных вариантов, для которых модель предсказывает изменение в функционировании клеток. Полученные результаты позволяют предполагать, что разработка может помочь врачам устанавливать причину наследственных заболеваний в большем количестве случаев и в перспективе решать ряд других важных задач. Например, создавать синтетические последовательности ДНК для включения генов только в одном или нескольких типах клеток человека – это может быть важным для генной терапии заболеваний человека.

Нейросеть разработана в рамках проекта «Cell State» по оценке роли мутаций в некодирующих белки регионах генома человека. Проект «Cell State» реализуется Институтом AIRI совместно с ИЦиГ СО РАН. Разработанная модель будет использоваться учеными для функциональной интерпретации генетических изменений в полногеномных данных.

Ознакомится с полным текстом препринта научной статьи можно по ссылке: <https://www.biorxiv.org/content/biorxiv/early/2022/01/01/2021.12.31.474623.full.pdf>

**Научно-Исследовательский Институт Искусственного Интеллекта AIRI** – автономная некоммерческая организация, занимающаяся фундаментальными и прикладными исследованиями в области искусственного интеллекта. Основная цель института - создание универсальных систем ИИ для решения задач реального мира. На сегодняшний день более 90 научных сотрудников задействовано в исследовательских проектах института для работы совместно с глобальным сообществом разработчиков, академическими и промышленными партнерами. АНО «Институт Искусственного Интеллекта» создана при поддержке Сбербанка.