

УДК 615.31, 004.032.26

## РАЗРАБОТКА ЛЕКАРСТВЕННЫХ СРЕДСТВ ДЛЯ ЛЕЧЕНИЯ АУТОСОМНО-ДОМИНАНТНОГО ПОЛИКИСТОЗА ПОЧЕК МЕТОДАМИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Капустина О.О. (Университет ИТМО), Серов Н.С. (Университет ИТМО)

Научный руководитель – д.х.н., профессор Виноградов В.В.  
(Университет ИТМО)

### Аннотация

В данной работе представлена новая концепция дизайна лекарств с использованием инструментов искусственного интеллекта. Универсальность концепции по отношению к различным заболеваниям в перспективе сможет изменить подход к разработке лекарственных средств.

### Введение

Разработка новых лекарственных препаратов – долгий, рутинный и дорогостоящий процесс. Из-за чего огромное количество редких и комплексных заболеваний остаются вне поля зрения фармацевтических компаний. Ярким представителем является поликистозная болезнь почек, приводящая к разрастанию кист в паренхиме почек и нарушению водно-солевого обмена в организме. Это наиболее распространенное генетическое заболевание, от которого, по данным Европейской ассоциации нефрологов, страдает более 12 млн человек. Однако терапевтический протокол уже многие годы не развивается и включает в себя только симптоматическое лечение, либо гемодиализ на терминальной стадии заболевания. Внедрение методов машинного обучения делает процесс создания лекарств более гибким, на три порядка дешевле и сокращает период разработки до 2-3 лет.

Существующие на данный момент экспериментальные препараты для лечения аутосомно-доминантного поликистоза почек (АДПП) не имеют доказанной высокой эффективности, поскольку воздействуют только на один из механизмов образования кист. Таким образом, перспективное соединение для лечения поликистоза должно одновременно воздействовать на ключевые мишени и быть безопасным при длительной терапии.

### Основная часть

Концепция нашего проекта заключается в применении генеративно-состязательных нейронных сетей для дизайна новой молекулярной структуры, потенциально обладающей фармакологической активностью. Архитектура модели включает в себя генератор и дискриминатор, в которой первый – генерирует набор данных, а второй – классифицирует образцы на реальные и поддельные. В процессе соревновательного обучения обе системы значительно повышают качество воспроизведения данных, вследствие чего сгенерированные образцы могут быть определены как настоящие. В настоящий момент алгоритм активно применяется для генерации изображений, однако создание молекул требует принципиально нового подхода, более тонкой настройки параметров и объема обучающих данных. В связи с этим, процесс генерации будет иметь три круга обучения. Первый будет реализовываться на обучающем наборе, состоящем из 11,5 тыс. молекул из базы данных PubChem и отсортированном в соответствии с критериями Липински. Так, модель научится распознавать общий паттерн для соединений, обладающих биологической активностью. Второй круг обучения включает в себя 5,5 тыс. молекул, одобренных Американской службой по контролю качества пищевых продуктов и лекарственных препаратов (Food and Drug Administration). Цель данного этапа – унаследование нейронной сетью ключевых особенностей, присущих лекарственным молекулам. На финальном этапе генерации задается набор молекул, которые обладают терапевтической активностью в отношении АДПП. Он состоит из 9 экспериментальных препаратов, находящихся на разных этапах исследований. Таким образом,

удастся создать набор уникальных молекулярных структур, с потенциальной активностью к целевому заболеванию.

Сгенерированные кандидаты являются исходным материалом, для дальнейшей разработки. Структуры необходимо проанализировать, исключить некорректные с точки зрения химического строения. Важным количественным показателем будет являться скоринг-функция – энергия, позволяющая оценить прочность связывания молекулы-лиганда с целевым белком, определение которой производится методами молекулярного докинга. Наиболее перспективными являются кандидаты, обладающие наибольшей аффинностью к одному или нескольким целевым белкам. Чтобы добиться специфической мультитаргетности молекул, конечная стадия концепции проекта включает в себя дополнительную оптимизацию молекулярной структуры кандидатов с применением генетического алгоритма. Тонкая настройка лиганда к мишеням сразу из двух групп, задействованных в механизме развития поликистоза значительно повышает вероятность высоких показателей клинической эффективности потенциального препарата на следующих этапах исследования.

На данный момент удалось построить архитектуру генеративно-состязательных нейронных сетей и запустить процесс обучения на первом, предварительно векторизованном, наборе данных. Мы столкнулись с ранним коллапсом модели, так как она очень чувствительна к изменениям гиперпараметров. Важно найти такие значения, при которых обе нейронных сети достигнут состояния равновесия и смогут воспроизводить качественные уникальные образцы молекул. Нами было замечено, что установление коэффициентов скорости обучения с разницей в 1,5–2 раз позволяют более точно корректировать веса на каждой итерации. Были получены более точные варианты изображений со значениями коэффициентов 0.0002 и 0.0004 для генератора и дискриминатора соответственно, однако требуется провести еще несколько циклов обучения с изменением параметров, чтобы сделать выводы об оптимальных величинах.

## **Выводы**

Текущие результаты указывают на перспективность реализации проекта в области разработки лекарственных средств. В будущем для нас важно добиться стабильности и качества задействованных алгоритмов и обеспечить глубокий теоретический анализ потенциального препарата.

## **Благодарности**

Работа выполнена при поддержке программы «Приоритет 2030»