



## Искусственный интеллект придёт на помощь разработчикам вакцин

Исследователи из группы "Биоинформатика" Института искусственного интеллекта AIRI представили SEMA (Spatial Epitope Modelling with Artificial Intelligence) – открытый инструмент для предсказания участков связывания антител с белками вирусов и бактерий. Инструмент поможет биологам и иммунологам в разработке вакцин и терапевтических препаратов на основе антител. SEMA позволяет предсказывать наиболее вероятные места посадки антител на исследуемый белок, например, S-белок вируса SARS-CoV-2, а также оценивать их эффективность для различных штаммов вируса. Совместно с учеными из НИЦЭМ им. Н. Ф. Гамалеи были проведены расчеты эффективности применения SEMA для коронавируса, а сейчас учеными проверяется возможность использования инструмента для вируса гриппа.

Основа инструмента – это нейросетевая языковая модель для предсказания эпитопов. Эпитопами называются привлекательные для антител участки на поверхности вируса, так называемые «посадочные площадки» для антител. Белки для анализа загружаются в программу в виде аминокислотной последовательности или 3D-структуры. SEMA способен предсказать, какие аминокислотные остатки этого белка формируют эпитопы, и оценить, насколько вероятно взаимодействие данного эпитопа с антителами. В дальнейшем эта информация может использоваться при выборе наиболее перспективных фрагментов белка для включения в состав вакцины, а также для поиска новых терапевтических антител против заданного участка белка. На независимом наборе тестов SEMA продемонстрировал более высокую точность по сравнению с другими известными инструментами.

Исследователи из AIRI продолжают работу над расширением функционала инструмента. В ближайшее время на ресурсе появится возможность предсказывать участки гликозилирования белков, то есть, определять места на поверхности белка, где расположены гликаны. Сами по себе гликаны – это покрывающие белки полимеры, и в норме они призваны защищать от вредных воздействий, однако, в ситуации, когда необходимо эффективное взаимодействие вирусного белка и антитела, они могут мешать реализации этого контакта.

В последнее время много исследований проводится с использованием методов обработки естественного языка применительно к последовательностям аминокислот в белках и нуклеотидов в геноме. Определение эпитопов на поверхности белков – задача не новая, однако, в SEMA исследователи впервые использовали для ее решения белковую языковую модель, а для удобства биологов создали веб-интерфейс. Вкупе с удобной визуализацией он призван упростить доступ ученых к сложным биоинформатическим инструментам на основе технологий искусственного интеллекта.

*«На сегодняшний день разработан ряд разрозненных инструментов для изучения белков в контексте их взаимодействия с антителами. Получение и объединение результатов работы с ними представляет большую сложность и требует много времени и ресурсов. Чтобы коллеги не тратили свое время на изучение кода, а сконцентрировались на биологической сути полученных*

*предсказаний и решении о дальнейшем их использовании в практике, мы создали веб-интерфейс», – старший научный сотрудник AIRI Татьяна Шашкова.*

Руководитель научной группы «Биоинформатика» Института AIRI **Ольга Кардымон** отмечает, что разработка инструмента велась в тесном контакте с учеными из Центра имени Гамалеи, поэтому коллектив надеется, что SEMA будет полезен и другим биотехнологическим институтам и компаниям, работающим над созданием вакцин и препаратов для терапии вирусных и бактериальных заболеваний.

По словам **Дмитрия Щербинина**, кандидата биологических наук, научного сотрудника лаборатории молекулярной биотехнологии и одного из разработчиков вакцины Спутник V, с помощью SEMA можно быстро просканировать несколько новых антигенных последовательностей и оценить угрозу ухода от иммунитета социально-значимых вариантов вирусов, например, новых штаммов вируса гриппа или нового коронавируса. *«Инструмент позволяет оценить наиболее опасные штаммы, сконцентрировать на них внимание и, таким образом, принять решение о своевременном реагировании на новые угрозы», – поясняет Дмитрий.*

*«Изучение перекрестного иммунитета – одна из важнейших задач современной вирусологии и иммунологии. Понимание антигенных взаимоотношений между циркулирующими штаммами позволяет использовать полученные знания для разработки эффективных вакцин и других препаратов; прогнозирования эпидемического процесса с более высокой точностью и принятия взвешенных решений по ограничительным мероприятиям. Наличие объемных баз данных, содержащих информацию о последовательностях различных вирусов, делает возможным применение искусственного интеллекта в научных исследованиях. При этом, в некоторых случаях основанные на ИИ решения могут показывать лучший результат, чем традиционные методы. На мой взгляд, SEMA именно такой случай», – Марина Асатрян, кандидат медицинских наук, старший научный сотрудник НИЦЭМ им. Н.Ф. Гамалеи.*

SEMA выложен в открытый доступ по ссылке <http://sema.airi.net>, разработчики могут найти исходный код здесь: <https://github.com/AIRI-Institute/SEMAi>. Результаты работы опубликованы в авторитетном научном журнале Frontiers in Immunology в секции Vaccines and Molecular Therapeutics (Q1).

**Научно-исследовательский Институт искусственного интеллекта AIRI** — автономная некоммерческая организация, занимающаяся фундаментальными и прикладными исследованиями в области искусственного интеллекта. На сегодняшний день более 90 научных сотрудников AIRI задействовано в исследовательских проектах Института для работы совместно с глобальным сообществом разработчиков, академическими и индустриальными партнерами