Искусственный интеллект, математика и физика

д.ф.-м.н., акад. РАН **Александр Михайлович Сергеев**

Национальный центр физики и математики, Саров;
Государственная корпорация по атомной энергии «Росатом», Москва

Использование искусственных нейронных сетей (ИНС) и глубокого машинного обучения для получения нового знания в математике и естественных науках привлекает в последнее время огромное внимание научного сообщества. Не вызывает сомнения, что ИИ в формате передовых больших языковых моделей превратился из удобного помощника в равноправного (по крайней мере) партнера в научном поиске, и темп развития этих моделей настолько стремителен, что заставляет задуматься о грядущей трансформации самой профессии ученого.

Алгоритмы ИНС прошли в своем развитии большой путь от перцептрона Розенблатта 1957 г. до трансформера компании Google 2017 г. и больших языковых моделей, от первых имитаций работы отдельных нейронов мозга до сложнейших архитектур, возможно, превосходящих по сложности коннектомы мозга высокоорганизованных существ. При этом машинные алгоритмы базируются на совершенно строгих математических теоремах, таких как теорема универсальной аппроксимации Цыбенко или теорема Колмогорова-Арнольда о представлении функции нескольких переменных через суперпозицию функций одного переменного.

Удивительным открытием последних месяцев стало то, что передовые ИНС класса ClaudeFable или GPTAstra, используемые математиками и физиками в режиме рассуждений, научились доказывать сложнейшие математические теоремы и находить новые аналитические решения базовых уравнений нелинейной физики. Примерами этого являются решение «задачи миллениума» о коллапсе исходно гладких распределений поля скоростей в трехмерном уравнении Навье-Стокса для несжимаемой жидкости и получение решений в магнитной гидродинамике для стационарных не тороидально-симметричных распределений плазмы в токамаках и стеллараторах (в опровержение так называемой гипотезы Грэда).

Помимо алгоритмической части в современных ИНС важнейшим вопросом является их аппаратная реализация на суперкомпьютерах с гетерогенной архитектурой. Помимо графических ускорителей передний фронт исследований здесь находится в области вычислений на новых физических принципах для увеличения скорости и снижения энергопотребления при работе с большими языковыми моделями. Недавние успехи в освоении фотонных вычислительных устройств позволяют рассчитывать на достижение в ближайшие годы скоростей обработки информации в системах ИИ до 10^{21} – 10^{22} операций в секунду, что на 2-3 порядка выше рекордной производительности сегодняшних ЭВМ.