



Представлен первый в мире публичный датасет для обучения ИИ самостоятельному принятию решений. Он состоит из 100 миллиардов фиксированных наборов данных

Контекстное обучение с подкреплением (In-Context Reinforcement Learning) – новое направление в исследовании ИИ, сфокусированное на создании моделей, которые не просто хорошо работают на всех тренировочных задачах одновременно, но и способны максимально эффективно дообучаться на новых за минимальное время. Например, большинство чат-ботов можно научить новой задаче, добавив в промпт некоторое количество примеров её решения. Во многих случаях это сработает, несмотря на то что при тренировке модели таким навыкам ее никто не обучал. Применение контекстного обучения с подкреплением упрощает работу создателей ИИ-моделей и позволяет облегчить выбор алгоритма, подбор гиперпараметров и имплементацию.

Наиболее успешно применять такие методы можно для того, чтобы обучить ИИ принимать решения. ИИ-агент должен исследовать (exploration) среду и понять, какие действия ведут к хорошему исходу, а какие к – плохому. Затем он должен остановиться и перейти в фазу эксплуатации хороших действий (exploitation). Задача exploration- exploitation заключается в том, чтобы найти механизм обучения, который на горизонте N шагов будет давать суммарно наилучшую награду, то есть получать наилучший результат за меньшее количество шагов.

В области ICRL нет стандартизации, и авторы всех ранее существующих работ были вынуждены каждый раз собирать новый датасет с нуля. Представленный научной группой «Адаптивные агенты» AIRI при участии лаборатории T-Bank AI Research датасет под названием XLand-MiniGrid доступен для всех ученых и разработчиков и насчитывает 100 миллиардов кортежей (фиксированных наборов данных). Он может пригодиться не только для создания ИИ-систем, которые будет гораздо проще настраивать для решения новых задач без вовлечения узкопрофильных специалистов, но и стать источником полезных синтетических данных.

Датасет был валидирован с помощью графиков обучающих историй – траектории были естественным образом упорядочены по возрастанию суммарной награды за эпизод.

«В свое время похожий на появление контекстного обучения с подкреплением прорыв произошел в области обработки естественного языка. Раньше люди обучали разные модели под разные задачи, но с появлением GPT-like методов смогли обходиться одной. Теперь вместо длительного процесса обучения для многих задач применяют лишь немного промпт-инжиниринга. Проблема в том, что GPT-like методы очень плохо решают задачи exploration-exploitation из-за специфики их данных – условно, весь интернет. Именно для этого и нужны такие датасеты. В будущем он может быть включен и в мультимодальные модели»

Владислав Куренков, руководитель группы «Адаптивные агенты» Института искусственного интеллекта AIRI

Получить доступ к датасету можно по ссылке: <https://github.com/AIRI-Institute/xland-minigrid-datasets>. Авторы работают над созданием научной публикации, посвященной процессу сбора и проверки датасета, а также развивают собственную среду для мета-обучения.

Вопросы: 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.