УДК 004

ИСПОЛЬЗОВАНИЕ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ В ЗАДАЧАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Хайруллин А.М.
ФГБОУ ВО «КГЭУ» г. Казань, Россия
Науч. рук. канд. тех. наук, доц. Р.С. Зарипова

Обучение с подкреплением – это одна из областей машинного обучения. Основная задача – обучение моделей для принятия последовательности решений. Модель учится достигать цели в неопределенной, потенциально сложной обстановке. Компьютер использует метод проб и ошибок, чтобы найти решение проблемы. Чтобы заставить машину делать то, что хочет программист, искусственный интеллект получает либо вознаграждение, либо наказание за действия, которые он выполняет. Его целью является максимизация общего вознаграждения.

Несмотря на то, что проектировщик задает политику вознаграждений – то есть, правила нахождения в определенной ситуации – он не дает модели никаких подсказок или предложений, как решить проблему. От модели зависит, как выполнить задание, чтобы максимизировать награду, начиная с совершенно случайных испытаний и заканчивая сложной тактикой. Используя возможности поиска и множества испытаний, в настоящее время наиболее эффективным способом намекать на креативность машины является подкрепление. В отличие от человека, искусственный интеллект может накапливать опыт из тысяч параллельных процессов, если алгоритм обучения с помощью подкрепления запущен на достаточно мощной компьютерной инфраструктуре.

Обучение моделей, управляющих беспилотными автомобилями, является прекрасным примером потенциального применения обучения с подкреплением. В идеальной ситуации компьютер не должен получать никаких инструкций по управлению автомобилем. Программист избегал бы точного описания всего, что связано с поставленной задачей, и позволял бы машине учиться на собственных ошибках. В идеальной ситуации единственным жестко закрепленным элементом будет функция вознаграждения.

Главная проблема в обучении с помощью подкрепления заключается в подготовке среды моделирования, которая в значительной степени

зависит от решаемой задачи. Когда модель должна стать сильнее человеческого интеллекта в, например, шахматах, то подготовка среды симуляции относительно проста. Когда речь заходит о создании модели, способной управлять автомобилем, очень важно построить реалистичный симулятор, прежде чем позволить машине ехать по улице. Модель должна придумать, как затормозить или избежать столкновения в безопасной среде.

Существуют и другие проблемы обучения с подкреплением, например, чрезмерно большая выборка данных, при которой возникает катастрофическое забывание старых знаний из-за получения новых. Или недостижение локального оптимума, при котором проблема решается неоптимальным путем. На сегодняшний день построенные модели могут быть далеко неидеальными, однако в недалеком будущем прогресс в этой области и движение к абсолютной точности неизбежны.

Источники

Килин Г.А., Ждановский Е.О. Преимущества использования обучения с подкреплением для обучения нейронных сетей // Материалы всероссийской научно-технической конференции: в 2х томах. 2018. С. 152-158.

Смирнитская И.А., Фролов А.А., Мержанова Г.Х. Модель выбора вознаграждения на основе теории обучения по подкреплению // Журнал высшей нервной деятельности им. И.П. Павлова. 2007. №2. С. 133-143.

Немчанинов А.В., Павловский Е.Н. Разработка метода разметки награды в обучении с подкреплением // МНСК-2018: Информационные технологии. Материалы 56-й Международной научной студенческой конференции. 2018. С. 241.

Научный руководитель  Зарипова Р.С.