

Топ 5 самых удивительных достижений машинного обучения за 2018 год

ПОДГОТОВИЛА УЧЕНИЦА 8А КЛАССА
СОГБОУ «ЛИЦЕЙ ИМЕНИ КИРИЛЛА И
МЕФОДИЯ»

МОЗГОВАЯ СТЕФАНИЯ

№5 Google neural machine translation

- ▶ Нейронная модель машинного перевода использует иные принципы работы с текстом, чем стандартный статистический метод перевода. До появления нейронных сетей перевод проводился пословно — система переводила отдельные слова и фразы с учётом грамматики. Поэтому при сложных оборотах или длинных предложениях качество перевода оставляло желать лучшего.
- ▶ GNMT же переводит предложение целиком, учитывая контекст. Система не запоминает сотни вариантов перевода фраз — она оперирует семантикой текста.
- ▶ При переводе предложение разбивается на словарные сегменты. Затем с помощью специальных декодеров система определяет «вес» каждого сегмента в тексте. Далее вычисляются максимально вероятные значения и перевод сегментов. Последний этап — соединить переведённые сегменты с учётом грамматики.
- ▶ А если учесть, что Google Translate работает с 10 000 языковых пар (103 языка), то понятно, что адекватные итоги можно подводить только сейчас.
- ▶ В ноябре 2016 года после полного окончания обучения системы и официального запуска аналитики Google представили [подробный анализ результатов GNMT](#). Из него следует, что улучшения в точности перевода несущественные — в среднем 10%.
- ▶ Наибольший прирост дали самые популярные языковые пары вроде испанский-английский или французский-английский — с результатом точности в 85–87%.

№4 Deepmind lip reading

- ▶ Не так давно ИИ приобрел еще один новый навык: чтение по губам. Правда, в данном случае без участия человека обойтись не удалось. Ученым пришлось установить в нейросеть продвинутый алгоритм распознавания мимики, после чего DeepMind предстояло «посмотреть» 5 тысяч часов шоу BBC, таких как BBC Breakfast, Newsnight и Question Time, в которых содержалось около 120 тысяч предложений. В результате этого ИИ научился «читать по губам» даже в тех случаях, когда человек «проглатывает» части слов. После этого началась самая интересная фаза эксперимента. Специалисту по распознаванию речи и искусственному разуму дали возможность посмотреть 200 не связанных между собой фрагментов. В итоге профессионал верно распознал лишь 12,5% слов, в то время как ИИ не ошибся в 46,8% случаев.
- ▶ Такого эффекта удалось добиться благодаря тому, что в процессе обучения алгоритм распознавания мимики, на который опиралась машина, был основан на достаточно продвинутой системе чтения по губам под названием LipNet, разработанной в Оксфордском университете. В данный момент специалисты из Google совместно с разработчиками LipNet готовят интернет-площадку со свободным доступом для разработчиков, занимающихся созданием систем автоматического чтения по губам для того, чтобы были созданы еще более совершенные механизмы распознавания. Как утверждает один из авторов эксперимента доктор Женьг Жоу,

№3 Deepmind Relation Network

- ▶ Реляционный модуль позволяет ИИ анализировать пары объектов и выводить соотношения между ними. К примеру, есть такой тип задач, когда нейронная сеть должна по фотографии ответить на какой-то вопрос к примеру: «Есть ли на картинке резиновая вещь того же размера, что и желтый металлический цилиндр?» До недавнего времени задача решалась с точностью всего лишь 68,5%, но нейросеть от deepmind с реляционным модулем достигла человеческой точности в 95,5%. Такая система может значительно улучшить алгоритмы визуального обучения, а также искусственный интеллект в виртуальных помощниках. Это технология на 3 месте, так как применение именно этой технологии довольно ограничено.

№2 Generative Adversarial Network

- ▶ GAN генерирует состязательные сети, такие как генератор и дискриминант. Во время обучения генератор из случайного вектора создает изображения и подает на вход дискриминатору, который пытается понять, реальная это картинка или фальшивка. Обычно они создают изображения спален или лиц, но в этом году они научились генерировать картинки по тексту. Изображение настолько реалистичное, что сложно поверить в то, что они сгенерированы по текстовому описанию. С этой технологией будет гораздо проще составлять фотороботы или работать с дизайнами, ну и наконец-то заказчики будут получать в точности то, что заказывали.

№1 Reinforcement learning

- ▶ Обучение с подкреплением. Суть в том, что сеть запоминает успешное поведение своего агента в заданной среде и получает при взаимодействии обратную связь. Нейросеть набирает знания через опыт, так же как и мы учимся в течении всей жизни. RL активно применяют в играх, роботах и в управлении трафиком. А вот чтобы научит робота ходить, прыгать, да еще и так, как человек, нужно сильно заморочится. Но исследователи все таки научили агентов эмулятора тел совершать пробежки с помощью конструирования сложной среды с препятствиями и с простым вознаграждением за прогресс в передвижении. Этот метод сокращает время обучения на 53%. Кстати, эта технология использовалась при тренировках альфа гоу от deepmind. Результаты и перспективность этой работы очевидны.



▶ Спасибо за внимание!