

Конспект лекции Николенко С.И.

1. История развития искусственного интеллекта

Развитие ИИ шло волнами. В 1950-е годы появились первые успехи: были созданы первые программы, способные решать логические задачи и играть в игры.

Затем наступили периоды спада, которые называют «зимами ИИ». В это время финансирование и интерес к области резко снижались из-за завышенных ожиданий и отсутствия реальных достижений. В середине 2000-х произошёл прорыв: нейросети наконец-то заработали! Началось бурное развитие глубокого обучения, что позволило решать задачи, которые раньше считались невозможными.

2. Современная ситуация и ускорение прогресса

После запуска ChatGPT (конец 2022 — начало 2023) темпы развития ИИ резко ускорились.

Сейчас важные события в области ИИ происходят не раз в несколько лет, а иногда даже несколько раз в день. Каждую неделю появляются новости, которые раньше становились бы главными событиями года или десятилетия. Прогресс не просто продолжается, а он ускоряется. Мы живем не в том таймлайне, где важные события происходят раз в пару лет, или даже реже, а где важные события могут происходить на 3 штуки в один день.

3. Возможность технологической сингулярности

ИИ становится всё более универсальным. Уже сейчас существуют модели, которые могут выполнять широкий спектр задач, ранее доступных только человеку. В будущем возможен переход к такому ИИ, который сможет делать всё, что умеет человек, но намного быстрее и без усталости. Это ведёт к концепции технологической сингулярности — моменту, когда ИИ начнёт сам себя улучшать, а люди потеряют возможность следить за этим процессом.

Мы стоим на пороге создания вообще искусственной цели. Такие AI-модели, которые будут способны делать все то, что мы с вами можем делать. Они будут это делать, как минимум, 24 часа в день. Когда искусственный интеллект начинает улучшать сам себя, потом еще и ещё, и ещё, то люди теряют возможность следить за этим прогрессом.

4. Риски и опасения

Эксперты и философы всерьёз обеспокоены рисками сверхинтеллекта. После появления ChatGPT опасения стали предметом публичных обсуждений. Если сверхинтеллект появится в ближайшие 5 лет, то нельзя полагаться только на философов — это может быть всего лишь промежуточный этап в эволюции интеллекта.

Сложности в управлении такими системами:

1. Чем мощнее становится ИИ, тем труднее его контролировать.
2. Возможна утрата контроля над развитием технологий.

Проблема рисков становится глобальным приоритетом. Вопросы безопасности ИИ обсуждаются на уровне государств, организуются международные саммиты, разрабатываются меры регулирования.

Проблема риска от искусственного интеллекта должна стать глобальным приоритетом, вместе с такими рисками, как пандемия и медленная война.

5. Позитивные стороны и польза ИИ

1. ИИ действительно меняет жизнь людей к лучшему.
2. Автоматизация рутинных задач, помощь в медицине, науке, образовании.
3. Возможность создавать новые продукты и услуги, ранее невозможные.

«Важно не останавливать прогресс, а сделать его безопасным и полезным для всех.» - Николенко С.И.

Рекомендуется ознакомиться с эссе о позитивных сценариях развития ИИ:

1. «Периферичный стейдж» (Семалькин)
2. «Мужчина со спаленной грейсисом» (Дарья Амадея)

«Искусственный интеллект, он все-таки действительно меняет жизнь людей к лучшему. И, конечно, нам не хочется все защищать и останавливаться. Нам хочется сделать такой прогресс, чтобы его пережить впоследствии. Жить в мире, который видит Дарья Матреть, очень хочется, и поэтому очень хочется все-таки исследоваться.» - Николенко С.И.

6. Прогнозы и законы развития

Прогнозы развития ИИ строятся *по экспоненте*. Часто используют аналогии с законом Мура, который описывает экспоненциальный рост вычислительных мощностей. С 2010-х годов темпы роста ИИ опережают рост самих вычислительных мощностей. Обучение новых моделей требует всё больше ресурсов (данных, вычислений, энергии). Открыты зависимости между размером модели и качеством её работы, которые хорошо подтверждаются на практике уже 7 лет.

«Люди рисуют прямые линии графика, как правило, это прямые линии географических шкафов, то есть это какие-то экспоненциальные передачи. В хорошем обучении есть и другой вид преодоления на графиках. Это прямые линии, которые буквально связывают размер модели и присутствие которой в процессе ее обучения и данные, которые ей нужны.» - Николенко С.И.

Заключение

Человечество стоит на пороге грандиозных изменений из-за ИИ. Необходимо трезво оценивать риски и потенциальную пользу, чтобы сделать прогресс безопасным и максимально полезным для общества.