Разве мы просто роботы?

6 марта 2025

Заглянем в будущее, зная прошлое и настоящее.

Искусственный интеллект сегодня оценивают все подряд. Наша газета попросила уделить внимание этой теме академика РАН Игоря КАЛЯЕВА, специалиста в области многопроцессорных вычислительных и управляющих систем, то есть эксперта в ИИ. Он заместитель академика-секретаря Отделения энергетики, машиностроения, механики и процессов управления РАН, заслуженный деятель науки Российской Федерации.

– Игорь Анатольевич, как трансформировался ИИ в последние годы? Какие задачи он будет способен решать лет через 10-15?

– Сегодня ИИ - один из основных трендов мирового научно-технологического развития. Вы наверняка слышали, что буквально на следующий день после своей инаугурации президент Трамп объявил о вложении 500 миллиардов долларов в развитие инфраструктуры ИИ в США.

Тем не менее в планетарные лидеры в области ИИ, с моей точки зрения, сегодня выходит Китай. Доказательством того стала появившаяся недавно информация, что в КНР созданы большие языковые модели, которые значительно эффективнее широко известного американского ChatGPT, и это сходу обрушило биржевую стоимость акций американских компаний в области ИИ.

Что касается главных трендов развития самого искусственного интеллекта, то основное внимание сегодня концентрируют именно на создании так называемых больших языковых моделей.

Эта технология появилась сравнительно недавно: в 2019 году американская фирма Open AI создала чат-бот Chat GPT2 на базе архитектуры нейронной сети, реализующей универсальный вычислительный механизм, так называемый трансформер. Он принимает на вход набор последовательных данных (например, слов текста) и формирует на выходе другой набор данных, преобразованный по специальному многопараметрическому алгоритму.

При этом разработчики использовали очень оригинальный подход к обучению такой нейронной сети: они просто выкачали из популярного онлайн-сервиса Reddit все текстовые сообщения, имеющие более трех лайков, и с их помощью обучили нейронную сеть-трансформер.

Таких ссылок было более 8 миллионов, а «вес» скаченных текстов превысил 40 Гбайт. Для сравнения: все творения Шекспира «весят» всего 5,5 Мбайт, а если вы будете читать по странице текста в минуту 24 часа в сутки, то вам понадобится 40 лет, чтобы достичь уровня «образованности» GPT2. Притом число варьируемых параметров в модели GPT2 более 1,5 миллиарда.

В результате GPT2 научился выдавать связные тексты, очень похожие на человеческие.

Но еще более впечатляющий скачок произошел в 2020 году, когда вышла новая версия - GPT3, в модели которой было уже 175 миллиардов параметров, а для ее обучения использовали около 420 Гбайт текстов. В результате неожиданно для всех GPT3 стал решать арифметические задачи, выполнять внятные логические рассуждения и т. д.

Что касается прогнозов развития ИИ на ближайшее десятилетие, то я считаю, что будет происходить постепенный переход от сравнительно общих систем, таких как ChatGPT, к специализированным моделям, работающим в узкой предметной области. И начало такого перехода мы уже наблюдаем на примере китайских DeepSeek и Qwen 2.5-Max. Благодаря суженной специализации они дают более эффективное решение при существенно меньших тратах на обучение.

– Какие угрозы может в себе таить столь стремительное развитие ИИ?

– Повсеместное внедрение ИИ порождает целый ряд проблем, которые могут привести к опасным последствиям для всего человечества. Во-первых, люди все больше становятся зависимыми от компьютерных систем и возлагают на ИИ функции, которые до недавнего времени считались прерогативой естественного человеческого интеллекта. А если повсеместная автоматизация и механизация труда, развитие телекоммуникационных технологий привели к появлению угроз, связанных с гиподинамией, то массовое использование ИИ сопряжено с рисками «умственной гиподинамии», т. е. деградации естественного интеллекта - системообразующего фактора человеческой цивилизации.

Вторая проблема - насколько мы можем доверять ИИ решение тех или иных задач. Если речь идет о тактических решениях, не приводящих потенциально к катастрофическим последствиям, то это вполне допустимо - мы же доверяем навигатору в машине. Но если речь о стратегических решениях, то доверять их ИИ рискованно.

Хочу напомнить известный случай из нашей недалекой истории, когда 26 сентября 1983 года ложное срабатывание системы предупреждения о ракетном нападении чуть не привело к началу ядерной войны между СССР и США. В тот день советская спутниковая система обнаружения стартов ракет «Око» выдала сообщение о запуске нескольких ракет с континентальной части США.

И только благодаря тому, что оперативный дежурный советского командного пункта не увидел характерных для запуска ракет вспышек и шлейфов огня, а также заметил, что радарное предупреждение не подтверждало пуск, ответный запуск не состоялся.
Как потом выяснилось, причиной ложной тревоги стала засветка датчиков спутника. А если бы в то время существовал ИИ и ему бы доверили принятие решение об ответном ударе по США? Он бы действовал по жесткому регламенту, и нас с вами сейчас бы просто не было.
Поэтому, считаю, необходимо четко определить зоны ответственности между естественным и искусственным интеллектом, особенно в решениях, чреватых глобальными катастрофическими последствиями.
Еще одним очень опасным, с моей точки зрения, следствием повсеместного использования ИИ является возможность манипулирования с его помощью человеческим сознанием путем изготовления различных дипфейков и интернет-влияния на психику человека.
Все это потенциально порождает новый тип войн - ментальный. Если в классических битвах целью является уничтожение живой силы и инфраструктуры противника, то в ментальной войне цель - уничтожение его самосознания и исторической памяти.
Причем последствия ментальной войны проявляются не сразу, только через поколение, когда исправить уже что-либо невозможно.
Вообще, очень многое в работе ИИ зависит от того, кто и на каких данных его обучал. От этого зависит, скажем так, менталитет ИИ. Буквально на днях опубликовали результаты исследований политических взглядов различных систем ИИ. И оказалось, что большинство ИИ поддерживает… левых политиков. Ведь основную массу моделей обучали на материалах и медиа, авторы которых тяготели к левым взглядам.
В газете The Washington Post Сэм Альтман, ведущий американский специалист в области ИИ, соучредитель и гендиректор OpenAI (создала ChatGPT), опубликовал статью, в которой утверждает, что только созданный США и их союзниками ИИ является «демократическим» и именно его надо использовать для внедрения западной этики и ценностей.
Все это в дальнейшем, видимо, приведет к разделению ИИ по политическому, национальному, половому, конфессиональному и прочим признакам. А как свидетельствует история человечества, именно такое размежевание - главная причина возникновения всевозможных конфликтов, вплоть до самых кровавых. В результате могут вспыхнуть войны между ИИ, в которых человечеству отведут роль малозначимого расходного материала.
– Так где ИИ все же стоит применять?
– Технологий ИИ уже сегодня находят широкое применение в самых различных сферах человеческой жизни и деятельности.
В промышленности - при прогнозировании спроса, оптимизации логистики поставок комплектующих, планировании производства.
В транспортной отрасли - при организации и налаживании грузоперевозок, создании беспилотного транспорта, обеспечении водителей средствами виртуальной и дополненной реальности с целью снижения аварийности.
В энергетике - при прогнозировании сетевых перегрузок и перераспределении мощности, предупреждении сбоя оборудования и систем.
В сельском хозяйстве - при роботизации сбора, переработки и транспортировки продукции, для дистанционного мониторинга и анализа состояния сельхозугодий.
В здравоохранении ИИ поможет переходу к персонализированной медицине, при прогнозировании распространения эпидемий и оценке генетических рисков. В банковском секторе - при прогнозировании динамики финансовых рынков.
В городской инфраструктуре - для оптимизации транспортных потоков, повышения качества обслуживания населения. В области безопасности и противодействия терроризму - при создании систем обнаружения лиц, находящихся в розыске, выявлении и предупреждении асоциального поведения людей, превентивного обнаружения террористических угроз в местах массового скопления народа.
Но наиболее значимый эффект сейчас ИИ дает на поле боя. Современную войну уже нельзя представить без ИИ, который используется при планировании военных операций, принятии тактических и стратегических решений, обработке разведданных, мониторинге угроз, распознавании целей, наведении на них средств поражения.
Одним из наиболее актуальных примеров применения такого ИИ является БПЛА. Современные беспилотники могут самостоятельно обнаруживать и наводиться на цели, работать в стае, выбирать оптимальные маршруты в обход средств РЭБ, противодействовать вражеским БПЛА и т. д.
Однако применение ИИ на поле боя ведет и ко все бÓльшим человеческим жертвам. Видимо, в будущем по аналогии с конвенциями о запрещении использования химического или ядерного оружия придется принимать какие-то международные законы, ограничивающие возможности ИИ в военных конфликтах.
– Какие важнейшие проблемы необходимо решить в области разработок ИИ в России?
– Самый важный вопрос. Для создания конкурентоспособного на мировых рынках ИИ нам нужны, во-первых, ученые, способные разрабатывать оригинальные модели ИИ, во-вторых, высококлассные программисты, реализующие процедуры машинного обучения этих моделей, и, наконец в–третьих, суперкомпьютерная инфраструктура, на основе которой эти модели ИИ должны физически существовать.

Первые две компоненты у нас есть, а вот с третьей проблема, хотя именно от нее прежде всего зависит успешное развитие ИИ. Ведь не зря Трамп решил потратить сотни миллиардов долларов на развитие суперкомпьютерной инфраструктуры ИИ.
Нам тоже необходимо создавать специальную суперкомпьютерную инфраструктуру для ИИ и машинного обучения на основе больших данных. Конечно, мы вряд ли сможем так просто, как американцы, взять и выделить 500 миллиардов долларов на это из бюджета страны, поэтому нам надо искать какие-то другие пути решения этой проблемы.
С моей точки зрения, данная проблема может быть решена путем организации Национальной суперкомпьютерной инфраструктуры (НСКИ), объединяющей в единую сеть все российские суперкомпьютеры, что обеспечит, во-первых, их широкую доступность для исследований и разработок в области ИИ, а во-вторых, возможность задействования всех суперкомпьютерных мощностей страны для реализации в том числе и больших языковых моделей ИИ.
Обидно то, что концепция создания такой НСКИ была разработана нашими ведущими учеными еще в 2019 году, прошла всестороннее рассмотрение и была одобрена Советом по приоритетному направлению Стратегии научно-технологического развития. Но с тех пор документ «ходит» по чиновничьим инстанциям, а за рубежом (например, в Китае) такая суперкомпьютерная инфраструктура активно создается. И результат налицо: китайские LLM уже успешно обгоняют американские.
– Вы наверняка слышали о машинном прототипе мозга. Он реально создается?
– С моей точки зрения, о каком-либо машинном прототипе мозга человека говорить рано. Пока ИИ физически живет на базе монструозных суперкомпьютерных систем, потребляющих десятки мегаватт энергии и имеющих гигантские габариты, в то время как мозгу человека для активности хватает всего 20 ватт и занимает он объем около 1400 см3.
Поэтому все наши потуги соорудить аналог человеческого мозга с помощью классических компьютерных технологий, по-моему мнению, обречены на провал. Последнее время все больше исследований посвящено разработке нейроморфных вычислительных систем, реализующих принципы обработки информации, присущие человеческому мозгу.
В частности, один из подходов связан с использованием мемристоров, то есть резисторов с эффектом памяти: их сопротивление зависит от прошедшего через них электрического заряда. Это свойство мемристоров напоминает свойства синапсов нейронов головного мозга, что, в свою очередь, открывает возможности эффективной реализации на базе мемристорных кроссбаров искусственных нейронных сетей.
Кроме того, поскольку мемристоры и живые нейроны работают на схожих принципах, то теоретически они могут «понимать» друг друга, что ведет к созданию нейрогибридных устройств, объединяющих в единое целое естественные и искусственные нейронные сети. А это, в свою очередь, - к нейропротезированию, т. е. замене естественных нейронных сетей, по каким-то причинам переставших выполнять свои функции, на искусственные нейронные сети, выросшие на базе мемристоров. Нужно только предварительно верно обучить такую искусственную нейронную сеть выполнению функций, которые реализовывала отмершая естественная нейронная сеть.
– Не опасна такая «игра в Бога»?
– Честно говоря, я не верю, что в обозримом будущем нам удастся создать ИИ, сравнимый с естественным человеческим интеллектом. Прошло более 70 лет с момента начала эры ИИ, а его цель - реализация искусственного аналога человеческого мозга - так же далека от нас, как и в начале пути. Недавно Илон Маск заявил, что в мире уже закончились реальные данные для обучения нейронных сетей, причем произошло это еще в 2024-м. И, несмотря на то, что ИИ «впитал в себя» все знания, накопленные людьми, он до сих пор не в силах достичь уровня естественного человеческого интеллекта. Это свидетельствует о том, что принципы работы естественного интеллекта кардинально отличны и, главное, гораздо эффективнее принципов работы современного искусственного интеллекта.
Нобелевский лауреат по физике Роджер Пенроуз в своей книге «Новый ум короля» утверждает, что человеческое сознание не является алгоритмическим, а его функции завязаны на некие квантовые эффекты, в силу этого оно не может быть смоделировано силами обычного компьютера типа машины Тьюринга.
В этом я с ним полностью согласен, а те квантовые эффекты в человеческом мозге, о которых говорит Пенроуз и которые мы до сих пор толком не понимаем, может быть, как раз и есть та самая «душа», без которой, никакое «железо» никогда по-настоящему мыслить и тем более обладать собственным сознанием не сможет.
Если же мы примем, что наш мозг алгоритмический, т. е. работающий по принципам обычного компьютера, то тогда нам также следует признать, что все мы с вами просто роботы, поскольку алгоритмический мозг (как и компьютер) может работать только по программе, заложенной извне.
Если же в будущем ИИ, превышающий по своим мыслительным возможностям естественный интеллект, создадут, боюсь, что того мира, в котором мы сейчас с вами живем, да и вообще всей человеческой цивилизации, скорее всего, просто не станет. Ведь Homo sapiens, т. е. человек разумный, в этом случае будет проигрывать естественный отбор ИИ и просто вымрет, как в свое время вымерли неандертальцы.
Но все-таки есть надежда, что этого не произойдет. По оценкам ученых, нашей Вселенной около 14 миллиардов лет. И по теории вероятности за это время где-нибудь в ней должна была возникнуть развитая цивилизация, которая гипотетически могла создать бесконечно развиваемый ИИ. И тогда такой

ИИ уже давно бы захватил и поработил всю Вселенную.
Но что-то и следов такого супер-ИИ мы нигде не наблюдаем. Скорее всего, это значит, что бесконечное и неконтролируемое развитие ИИ просто невозможно.

Елизавета ПОНАРИНА