

Для цитирования:

Рослякова Н.А. Искусственный интеллект: вопросы восприятия и включения в жизнь // Наука о данных : материалы международной научно-практической конференции. Санкт-Петербург, 5–7 февраля 2020 г. – СПб. : Изд-во СПбГЭУ, 2020. – С. 260-261.

Рослякова Н.А.

Искусственный интеллект: вопросы восприятия и включения в жизнь

Математика на протяжении всей истории науки определяла развитие и понимание природы разнообразных процессов. В целом переход от описательных к аналитическим возможностям науки неразрывно связан с математикой. В последние несколько десятилетий наблюдается «бум» развития нейросетей или искусственного интеллекта (ИИ), компьютерного зрения, распознавания текста, генерирования речи и т.п. При этом даже специалисты отдельных узких областей знаний утверждают, что не существует понимания многих принципов действия ИИ, и говорят о важности появления новой теории математики. Для обывателей широкое вхождение в жизнь технологий ИИ для решения частных задач часто становится базой для мифотворчества. Разнообразные мифы рождаются во многих областях и всегда они связаны с отсутствием понимания принципов действия различных технологий. Некоторые люди не используют микроволновые печи так как считают, что меняется молекулярная структура продуктов и происходит вредное облучение. Некоторые люди не используют рентген и флюорографию опасаясь большого дозы излучения. Таких примеров масса. Это происходит от непонимания и ошибочной интерпретации сложных технологий.

При этом внедрение новых технологий порождает развитие теории, которая позволяет более полно описать принципы действия технологии. Т. Сейновски приводит пример появления теории термодинамики как способа осмыслить и усовершенствоваться технологию парового двигателя (Возможен ли ..., 2019). В свою очередь новые теории становятся основой для переосмысления технологий и их дальнейшего развития (термодинамика в дальнейшем позволила сконструировать мощнейшие локомотивы).

Центральной темой современного научного дискурса является проблема неполной ясности принципов и закономерностей работы ИИ. Теоретические основы нейросетей, который сейчас испытывают «бум» развития, были заложены в 80-е гг. прошлого столетия. Разумеется, такое неполное знание обостряет дискуссию о том, нужно ли понимать принцип работы ИИ или возможно пользоваться результатами его работы и не вникать в то, как ИИ получил этот результат (то есть, воспринимать его как «чёрный ящик»). В данном вопросе эксперты сходятся во мнении, что понимание необходимо.

Проблемой, связанной с неполным знанием о технологиях (ИИ, в частности) являются различные сбои и катастрофы, которые возникают в процессе их использования. При развитии ИИ вероятны такие ситуации как Чернобыльская и Фукусимская аварии в атомной энергетике. В этом случае часто возникают предложения свернуть использование таких технологий, отказаться от разработок, что является в корне ошибочной позицией. В этом смысле развитие теории об ИИ является не только основой для будущего развития науки, но и способом усовершенствовать эту уже внедрённую технологию.

С точки зрения практического применения технологий и разработки конкретных продуктов на их основе критическое значение приобретает проблема обоснования решений ИИ. В самом простом варианте при построении автомобильного маршрута мы добавляем информацию о пробках, чтобы понять, почему маршрут приложен каким-то непривычным образом. В более критических случаях (например, когда ИИ в медицине предложит ампутировать пациенту ногу) нужно будет серьёзное обоснование этого решения, иначе согласиться с ним будет крайне сложно. В общем смысле такое понимание и объяснение работы ИИ позволяет создавать максимально качественный продукт.

ИИ уже сейчас решает отдельные задачи быстрее и успешнее, чем человек (компьютерное зрение, направленное на выявление мошенников, террористов), что позволяет получать положительные эффекты в краткосрочном горизонте. То есть, инвестирование в развитие и освоение технологий компьютерного зрения, интернета вещей, цифровых двойников, распознавания и синтеза речи обуславливает конкурентоспособность. В среднесрочной перспективе расширение использования и совершенствование технологий ИИ актуально для всей экономики (и финансово-сервисного сектора, и промышленных компаний): если не вкладывать в ИИ сейчас, то перспективе 5-10 лет компании безнадёжно отстанут.

С другой стороны, проблема ИИ и знаний о нём также выражается в том, что не всегда понятно для каких задач он успешен и применим, а для каких – нет. В частности, нет возможности обанкротить 100 тыс. банков для того, чтобы научить нейросеть управлять банком, поэтому необходимо пользоваться абсолютно понятными моделями, которые являются полностью интерпретируемыми. В свою очередь, возможно выделить задачи, которые являются очень хорошей зоной для осуществления различных экспериментов. В качестве примера можно привести рекомендации музыки или статей. В случае и удачного, и неудачного результата происходит обучение и совершенствование системы.

При этом у обывателя может возникнуть иллюзия, что ИИ может всё за них решить и отрегулировать. Преодоление этого заблуждения заключается в том, что человек должен стремиться к тому, чтобы понимать в каком мире он живёт. В этом смысле необходимы некоторые усилия государства для получения каждым

человеком понимания природы ИИ. В этом состоит основа следующего витка развития общества. Неразрывна с совершенствованием математической культуры и культур взаимодействия с ИИ проблема построения партнёрских отношений с ИИ. Она выражается в том, что многие виды деятельности видоизменяются и будут невозможны без ИИ. Самый эффективный доктор и ИИ действуя по отдельности имеют эффективность 92%, тогда как их партнёрство могло бы увеличить эффективность до 99% (Assessing the ..., 2018, с. 46). Это комплексная проблема, которая требует разносторонних подходов для решения.

Решения:

- развитие теорий обучения, повышение требований к обучающим выборкам позволит добиться, чтоб ИИ не только принимал решение, но и генерировал некоторое объяснение;

- краудсорсинг и инженерный подход при взаимодействии с ИИ могут повысить эффективности работы множества специалистов и общества в целом;

- массовая подготовка кадров и организация совместной работы с ИИ.

При этом важно отметить, что понимание ИИ со стороны специалистов и со стороны обывателей будет существенно отличаться. По нашему мнению, важность формирования именно обывательского, «человеческого» понимания природы ИИ заключается в том, чтобы научить общество относиться к ИИ как к инструменту и понимать его возможности и опасности. Ситуация, когда такое понимание не будет сформировано в широком масштабе, является серьёзной основой для усугубления каких-то нештатных ситуаций. То есть, в критический момент из-за отсутствия знания люди будут поступать наиболее опасным и вредным для себя способом. Для предотвращения подобных ситуаций при взаимодействии с ИИ, конечно, важно проводить ликбез в этой сфере. В частности, одним из важнейших направлений в будущем может быть формирование методов и инструментов проведения государственной информационной кампания по распространению знаний об ИИ. Могут быть оформлены сайты, социальная реклама, научно-популярных изданиях для старшего населения, проведены специальные тематические встречи в школах, вузах для более молодого поколения.

Это есть основа для формирования доверия к ИИ. В недалёком будущем 20-30 лет прогнозируется появление у ИИ здравого смысла, элементов рефлексии, задатков этики (прижимание о возможном). Всё это делает всех людей включёнными в процесс взаимодействия с ИИ и, следовательно, требуется подготовиться к этому, чтобы потом испытывать как можно меньше проблем.

Источники

Возможен ли (искусственный) интеллект без математики? / Фонд «Росконгресс». 2019. Режим доступа URL: <https://roscongress.org/sessions/spief-2019-vozmozhen-li-iskusstvennyy-intellekt-bez-matematiki/translation/> (дата обращения: 10.01.2020).

Assessing the Economic Impact of Artificial Intelligence / International Telecommunication Union. 2018. URL: https://www.itu.int/dms_pub/itu-s/opb/gen/S-GEN-ISSUEPAPER-2018-1-PDF-E.pdf (дата обращения: 10.01.2020).