21 октября 2021

Не в тех руках: как можно навредить с помощью ИИ

Светлана Вронская,
Эксперт департамента аналитических решений ГК "КОРУС Консалтинг"

Искусственный интеллект прочно вошел в нашу жизнь, и, бесспорно, заметно ее облегчил. Но не все стали использовать его в благих целях. Последние десять лет исследователи в разных странах изучают угрозы, которые злонамеренные использования ИИ представляет для общества в целом и отдельных сфер человеческой деятельности – в политике, экономике, оборонно-промышленном комплексе. Наш эксперт Светлана Вронская, не раз я рассказывала о пользе искусственного интеллекта для всех сфер – от производства и телекоммуникаций до маркетинга, видеонаблюдения и ценообразования. Сегодня поговорим об обратной стороне медали и рассмотрим, как ИИ может навредить человеку.

Но перед этим давайте договоримся об одном: искусственный интеллект не опасен. Это просто компьютерная программа. Она написана людьми,

управляется людьми и может быть уничтожена людьми. «Кого же нам тогда опасаться?» – спросите вы. А опасаться надо того, кто владеет технологиями разработки AI-решений. Именно такие люди, злонамеренно использующие или потенциально имеющие возможность использовать ИИ, заставляют исследователей делить риски на три группы. Рассмотрим каждую из них.

Кибербезопасность

Хакерские атаки, сливание в сеть персональных данных клиентов различных компаний, рассылка массовых вирусов – уже давно это не является для нас новостью. Однако, чем опасны в этом отношении системы на базе искусственного интеллекта, так это в том, что AI-злодеи способны на большее, чем просто залезть в вашу компьютерную сеть по предписанному алгоритму. Оказавшись внутри сети или устройства, эти злонамеренные системы «думают» о том, как нанести максимальный ущерб.

Один из ярких примеров в этой области показала недавно компания IBM. Корпорация создала систему для видеоконференцсвязи Deeplocker. Правда цель у этой программы была не совсем обычная. DeepLocker нес в себе заряд вируса WannaCry. Если бы этот вирус запустили во всю компьютерную сеть, то велика вероятность, что системные администраторы распознали бы угрозу. Затем они бы нейтрализовали ее с компьютеров тех пользователей, которые и были основной целью заражения и ничего плохого не случилось. Вирус уничтожили еще до того, как он дошел до них. Именно поэтому и придумали DeepLocker. Во время видеоконференции программа сканирует лица собеседников, находит и распознает с помощью встроенного ИИ-

инструмента нужного человека и в его компьютер подселяет вирус. Конечно, IBM придумал этот продукт в качестве теста-предостережения, но эксперимент заставляет задуматься.

Количество кибератак будет расти, а с ним участиться и использование в них искусственного интеллекта. Уже сейчас, согласно исследованию производителя антивирусов McAfee, каждую минуту совершается 375 попыток киберпреступлений.

Жизнедеятельность человека

Примером может служить ситуация, в которой злоумышленник захватывает системы управления беспилотным трафиком или энергетическими комплексами и намеренно вызывает аварии, взрывы и прочие катастрофы.

Специалисты прогнозируют, что защититься в этой области от искусственного интеллекта будет труднее всего, так как для этого потребуются инвестиции в оборудование и программное обеспечение, которые сможет предотвращать подобные роботизированные атаки.

Дестабилизация международных отношений

Звучит несколько пафосно, но суть проста. Давайте вообразим себе, что нам и нашим друзьям рассылают в мессенджере информацию о митинге на какую-то очень важную для нас тему. Мы приходим на этот митинг, а там... Дальше может быть что угодно. И мы никогда не узнаем, что сообщения рассылал чат-бот, чьей задача было сконструировать сообщение о подложном

мероприятии и убедить потенциальных жертв его посетить.

Менее страшные, но от этого не менее неприятные примеры – это сбор и обработка данных о пользователях соцсетей для запуска точечной пропаганды. Все помнят скандал с Cambridge Analytica, когда компанию обвинили в использовании личных данных 50 миллионов пользователей «Фейсбука» для того, чтобы повлиять на исход президентских выборов в США в 2016 году. Но ведь большая часть айсберга просто скрыта от публики.

Недавно американская технологическая компания NVIDIA поделилась результатами работы генеративно-конкурентной сети, обученной самостоятельно генерировать фотографии людей (fake people). Нейросеть за секунды создает в большом разрешении изображения лиц несуществующих людей, может добавлять любые культурные и этнические особенности, эмоции, настроение и основана на бесконечной коллекции картинок реальных лиц. А потом, потенциально, это изображение может обращаться лично к вам и что-то. И, поверьте, вам будет очень трудно отказаться, так как этот fake person будет сконструирован так, чтобы вызвать эмпатию именно у вас.

Как защититься от AI

Что же сделать, чтобы обезопасить себя и своих близких от подобных угроз? Обычно пишут очень общие вещи, сводящиеся к тому, что нужно жить дружно и соблюдать этический кодекс использования искусственного интеллекта. Я против неконкретных рекомендаций, поэтому делюсь своими мыслями о том, что же делать, чтобы ИИ-преступники не застали нас врасплох.

Во-первых, изучать и использовать искусственный интеллект. Согласно исследованию Capgemini Research Institute, 69% компаний в мире считают, что противостоять подобным угрозам без AI невозможно. Поэтому давайте будем на шаг (а лучше на два) впереди злоумышленников, создавая системы защиты еще до того, как они нам понадобятся.

Во-вторых, думать. Понимаю, что это звучит банально, но подумайте сейчас, скольких «друзей» в соцсетях вы знаете лично? Привыкайте общаться только с теми, кого вы персонально знаете и можете проверить поступившую информацию. Изучайте первоисточники новостей и находите подтверждение каждой самостоятельно.

В-третьих, использовать искусственный интеллект для благих целей. Существует целый пласт кейсов использования AI, не выходящих за рамки этики, которые, тем не менее вызывают нарекания. Могу привести в качестве примера так называемый predictive policing – применение ИИ в правоохранительных органах для предотвращения преступлений. В США эта система уже давно вызывает вопросы, так как склонна безапелляционно относить к группе подозреваемых чернокожих молодых людей, а в некоторых государствах от подобных систем уже отказываются.

Подобные истории можно встретить и с применением искусственного интеллекта в судебной системе, в системе образования (при процедуре поступления в ВУЗы) и других отраслях, где AI служит платформой для дискриминации людей по таким признакам, как пол, цвет кожи, возраст, доход и так далее.

***

Что хочется сказать в конце? Мой главный посыл на сегодня – вкладывать как можно больше ресурсов в разработку ИИ-решений для мирных целей. Чем больше мы этому уделим внимание сейчас, тем меньше времени у злоумышленников будет на запрещенные деяния.

Источник: VC.ru