

ЭТИЧЕСКИЕ ПОСЛЕДСТВИЯ И РЕШЕНИЕ ПРОБЛЕМ ИСПОЛЬЗОВАНИЯ ЯЗЫКОВЫХ МОДЕЛЕЙ В ПРОФИЛАКТИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ ПОЛИЦИИ

Аннотация: Правоохранительные органы должны установить этические руководящие принципы, регулирующие использование технологий прогнозирования деятельности полиции. Во избежание этических последствий, эти руководящие принципы должны способствовать смягчению предвзятости, конфиденциальности данных и соблюдению гражданских свобод, уменьшая риски и обеспечивая их положительный вклад в развитие общества.

Ключевые слова: искусственный интеллект, внедрение языковых моделей, этические последствия, прозрачность, подотчетность, конфиденциальность данных.

С развитием технологий искусственный интеллект [ИИ] был интегрирован в современную юридическую практику для оптимизации ее деятельности, повышения эффективности и улучшения процесса принятия решений, добившись значительных успехов в ее преобразовании. Инструменты и программное обеспечение на базе искусственного интеллекта были разработаны с целью оказания помощи правоохранителям в различных аспектах их работы, включая юридические исследования, анализ документов, управление контрактами и предиктивную аналитику [1].

Сегодня алгоритмы искусственного интеллекта могут анализировать огромные объемы данных и предоставлять информацию, на раскрытие которой людям требуется значительно больше времени. Системы на базе искусственного интеллекта могут проверять и анализировать контракты для выявления потенциальных рисков или несоответствий, что не только экономит время, но и снижает вероятность человеческой ошибки, т.е. исключает пресловутый 'человеческий фактор'. Более того, ИИ помогает автоматизировать рутинные задачи, подобных составлению документов или комплексной проверке, высвобождая время сотрудников с тем, чтобы они могли сосредоточиться на более сложных юридических вопросах.

Однако использование искусственного интеллекта в юридической отрасли вызывает беспокойство на предмет конфиденциальности и безопасности данных. Системы искусственного интеллекта полагаются на огромное количество конфиденциальной информации, включая личные данные клиентов, конфиденциальные документы и привилегированные сообщения. Крайне важно обеспечить наличие надлежащих мер безопасности для защиты этих данных от несанкционированного доступа или неправильного использования. Кроме того, существует риск того, что предвзятые или дискриминационные алгоритмы могут увековечить существующее неравенство в правовой системе [2].

Большая языковая модель – это модель искусственного интеллекта, которая обучается на огромном наборе данных, чтобы позволить конкретной системе ИИ понимать, переводить и генерировать человеческий язык. Эта модель важна в передовых приложениях обработки естественного языка или НЛП. Следует отметить, что существуют даже специальные системы искусственного интеллекта, основанные на продвинутых больших языковых моделях, которые могут обеспечивать человеческие ответы и полноценные человеческие разговоры.



Принимая во внимание все преимущества интеграции языковых моделей в предиктивную полицейскую деятельность, действительно способных произвести революцию в правоохранительной практике, важно учитывать проблемы и этические соображения для обеспечения ответственного и справедливого использования данной технологии.

Известно, что этика искусственного интеллекта выделяет следующие этические проблемы, возникающие при использовании технологий ИИ:

1) этические проблемы, связанные с системами искусственного интеллекта как объектами, то есть инструментами, созданными и используемыми людьми (конфиденциальность, непрозрачность, предвзятость);

2) этические проблемы, связанные с системами искусственного интеллекта как субъектами, то есть этика самих систем ИИ (искусственная мораль, машинная этика);

3) проблема возможного будущего сверхума искусственного интеллекта, приводящая к неконтролируемому и необратимому развитию искусственного интеллекта, что приведет к радикальному изменению природы человеческой цивилизации [3].

Следующие вопросы подчеркивают некоторые ключевые проблемы и дают представление о том, как их можно эффективно решить [4]:

Конфиденциальность и защита данных. Использование языковых моделей в профилактической полицейской деятельности требует доступа к огромным объемам данных, включая личную информацию. Крайне важно установить надежные меры защиты данных и соблюдать правила конфиденциальности для защиты прав людей. Должны существовать строгие протоколы, гарантирующие, что данные получены законным путем, надежно хранятся и используются только в законных целях правоохранительных органов.

Предвзятость и справедливость. Языковые модели обучаются на существующих данных, которые могут содержать предвзятость, свойственную системе уголовного правосудия. Крайне важно признавать и смягчать эти предубеждения, чтобы предотвратить несправедливое преследование или профилирование конкретных сообществ или отдельных лиц. Алгоритмы обнаружения систематических ошибок могут использоваться для выявления и устранения любых подобных систематических ошибок, присутствующих в обучающих данных. Регулярные аудиты и постоянный мониторинг результатов работы системы могут помочь выявить и исправить любые непреднамеренные отклонения, которые могут возникнуть во время ее использования [5].

Прозрачность и подотчетность. Для укрепления общественного доверия прозрачность и подотчетность должны быть на переднем крае внедрения языковых моделей предиктивной полицейской деятельности. Правоохранительные органы должны предоставить четкие инструкции о том, как используются языковые модели, какие данные собираются и как результаты технологии используются в процессах принятия решений. Кроме того, должны быть созданы механизмы, позволяющие проводить независимые проверки и надзор для обеспечения соблюдения этических стандартов и предотвращения злоупотреблений.

Человеческий надзор и суждение. Языковые модели не должны заменять человеческое суждение, а скорее служить инструментами для поддержки и улучшения процессов принятия решений. Эксперты-люди, такие как специалисты правоохранительных органов и юристы, должны участвовать в интерпретации результатов языковых моделей, проверке информации и принятии окончательных решений. Люди несут ответственность за оценку контекста, рассмотрение дополнительных доказательств и обеспечение соблюдения правовых и этических стандартов.

Постоянная оценка и совершенствование. Использование языковых моделей в предиктивной полицейской деятельности должно представлять собой итеративный процесс с постоянной оценкой и улучшением. Правоохранительные органы должны регулярно



оценивать эффективность и влияние языковых моделей, принимая во внимание отзывы заинтересованных сторон и вовлеченных сообществ. Этот процесс оценки должен включать механизмы выявления и исправления любых недостатков или непредвиденных последствий используемой технологии [6].

Вовлечение сообщества и сотрудничество. Вовлечение сообщества во внедрение и надзор за системами прогнозирования полицейской деятельности на основе языковых моделей имеет важное значение. Открытый диалог, инициативы по вовлечению сообщества и консультации с общественностью могут помочь решить проблемы, собрать информацию и гарантировать, что технология используется в соответствии с ценностями и приоритетами сообщества. Сотрудничество с правозащитными организациями, научными кругами и экспертами в этой области может предоставить ценную информацию и рекомендации на протяжении всего процесса. Такое сотрудничество имеет тенденцию открывать более широкие возможности для академических учреждений для проведения исследований и предоставления научно обоснованной информации об эффективности и этических последствиях внедрения языковых моделей.

Междисциплинарные исследования и инновации. Сотрудничество между научными кругами, разработчиками искусственного интеллекта и правоохранительными органами может стимулировать междисциплинарные исследования и инновации в области предиктивной деятельности полиции на основе языковых моделей. Исследователи могут изучить новые подходы для повышения точности и эффективности языковых моделей, а также решить проблемы, связанные с предвзятостью, объяснимостью и интерпретируемостью. Совместные усилия могут привести к разработке передовых методов и инструментов, которые еще больше улучшат системы прогнозирования полицейской деятельности на основе языковых моделей [7].

Совместные дискуссии между государственными органами, правоохранительными органами и экспертами по правовым вопросам имеют решающее значение для разработки политики и правовой базы, регулирующей использование языковых моделей в профилактической полицейской деятельности. Эти рамки должны учитывать права на конфиденциальность, защиту данных, прозрачность и подотчетность, а также решать конкретные проблемы и требования технологии языковых моделей. Совместные усилия могут привести к разработке передовых методов и инструментов, которые еще больше улучшат системы прогнозирования полицейской деятельности на основе языковых моделей и установят баланс между общественной безопасностью и защитой гражданских свобод [8], [9].

Активно решая эти проблемы и интегрируя этические соображения во внедрение систем прогнозирующей полицейской деятельности на основе языковых моделей, можно снизить риски и максимизировать преимущества этой технологии. Ответственное использование языковых моделей может помочь правоохранительным органам в их миссии по обеспечению общественной безопасности, одновременно соблюдая фундаментальные права и ценности [10].

Список литературы:

1. Harry Surden, Artificial Intelligence and Law: An Overview, 35 Ga. St. U. L. Rev. (2019). Available at: <https://readingroom.law.gsu.edu/gsulr/vol35/iss4/8>
2. Васильев, А.А., Печатнова, Ю.В. Искусственный интеллект и право: проблемы, перспективы / А.А. Васильев, Ю.В. Печатнова // Российско-азиатский правовой журнал. – 2020 – № 2. – С. 14-18.
3. Шляпников, В.В. Некоторые проблемы этики искусственного интеллекта / Б.Б. Шляпников // Идеи и идеалы. – 2023. – Т. 15, № 2, ч. 2. – С. 365–376.



4. Игнатьев, А.Г. Этика в области искусственного интеллекта в фокусе междисциплинарных исследований и развития национальных подходов /А.Г. Игнатьев. – М.: 2022. – 26 с.

5. Поляков, М.П., Смолин, А.Ю. Цифровые технологии на службе правосудия: идеологические предпосылки и технические перспективы / М.П. Поляков, А.Ю. Смолин // Юридическая наука и практика: Вестник Нижегородской академии МВД России. -2018. -№2. – 42 с.

6. Fastovich, G. G. On the issue of information technologies in modern Russia (on the example of the electronic document study management system) / G. G. Fastovich, S. A. Kapsargina, I. V. Kudashova // Евразийский юридический журнал. – 2021. – No. 11 (162). – P. 157-159.

7. Войниканис Е.А., Семенова Е.В., Тюляев Г.С. Искусственный интеллект и право: вызовы и возможности самообучающихся алгоритмов /Е.А. Войниканис, Е.В. Семенова, Г.С. Тюляев // Вестник ВГУ. Серия: Право. -2018. -№4 (35).

8. How Will Artificial Intelligence Affect Policing and Law Enforcement? – Artificial Intelligence + (aplusinfo.com)

9. Gotsko, L. G. Large language models (LLM) benefits and risks for ethical compliance / L. G. Gotsko //, 23 мая 2024 года, 2024. – P. 33-36.

10. Гоцко, Л. Г. Принципы этики в контуре новой технологической парадигмы развития искусственного интеллекта / Л. Г. Гоцко // Философия в XXI веке: направления и тенденции развития: Материалы II Международной научно-практической конференции: В 3-х частях, Москва, Зеленоград – Красноярск, 12 апреля 2024 года. – Москва: федеральное государственное автономное образовательное учреждение высшего образования "Национальный исследовательский университет "Московский институт электронной техники", 2024. – С. 69-78.

