

АНАЛИЗ АЛГОРИТМОВ ПРИНЯТИЯ РЕШЕНИЙ

Криштафович К.Д., Внук О.М., Плиска В.С.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Рогов М.Г. – ассистент кафедры информатики

Аннотация. Алгоритмы принятия решений – вычислительные методы, позволяющие машинам автоматически выполнять задачи и/или принимать решения – становятся ключевым фактором, обеспечивающим конкурентоспособность организации. Однако, несмотря на множество преимуществ, существует обратная сторона, связанная с их использованием, которая может привести к некоторым проблемам. Важно понимать потенциально негативные последствия, которые могут возникнуть в результате применения алгоритмов принятия решений, ещё до того, как они полностью проявятся. Целью данной работы является выявление, анализ и объяснение проблем, которые могут возникнуть, а также предоставление пятиэтапного плана по внедрению алгоритмов в работу.

Рассмотрим сценарии, при которых алгоритмы принятия решений могут влиять на ход жизни людей. Может ли получение страховки на жизнь зависеть от профиля в Instagram? Как будут себя чувствовать люди, если медицинская страховка будет требовать обязательного ношения фитнес-трекера для обеспечения отслеживания состояния. Для некоторых эти сценарии могут показаться нереальными, однако они уже являются частью реальности. Алгоритмы играют важную роль в принятии решений страховыми компаниями относительно качества обслуживания или отказа клиентам. В качестве примера можно привести компанию по автомобильному страхованию «American Automobile Association», которая предлагает приложение «Drive». Оно оценивает навыки вождения по шкале от 0 до 100. Если водитель соблюдает скоростной режим и в целом водит безопасно (это определяют датчики в автомобиле) ему предоставляются сниженные страховые взносы. Аналогично, компания по медицинскому страхованию «John Hancock» ввела программу «Vitality». После согласия предоставлять данные для анализа о физической активности с помощью фитнес-трекера клиенты могли либо снизить страховые взносы, либо. Кроме того, в 2014 году компания «State Farm» представила подход, который позволяет анализировать данные клиентов, включая информацию о доме, автомобиле и здоровье. Искусственный интеллект (ИИ), машинное обучение, а также алгоритмы, использующие эти технологии, часто представляются как одни из перспективных технологических достижений, доступных человечеству. Потенциальные преимущества ранжируются от более эффективных стратегий до улучшения обслуживания и увеличения выручки. Однако гораздо реже обсуждаются негативные последствия. Это является проблемой по двум причинам. Во-первых, возрастающая зависимость общества от алгоритмов, в сочетании с непредсказуемым будущим технологий, увеличивает вероятность возникновения негативных последствий. Во-вторых, без полного понимания того, как алгоритмы принимают решения и какие могут быть последствия от их использования, люди не смогут ответственно использовать эти технологии.

Алгоритмы могут принимать решения благодаря огромным объемам данных. Можно считать большие данные «цифровыми следами», полученными в результате сложных взаимодействий между людьми и машинами. Поэтому качество наборов данных, используемых для разработки или обучения алгоритмов, влияет на их возможности и последующие решения. Это приводит к трем основным проблемам. Во-первых, наборы данных часто обладают низким качеством. При создании «больших» наборов данных разные форматы объединяются из различных источников, что приводит к недостаточной актуальности или полноте полученных наборов. Известно, что данные, собранные фитнес-трекерами, все чаще оспариваются в способности точно измерять активность. Еще одной причиной плохого качества данных является отсутствие метаданных – важных контекстных сведений или «данных о данных». Во-вторых, алгоритмы имеют ограничения в своей способности принимать решения. На самом базовом уровне алгоритмы ищут зависимости в данных и принимают решения на основе них. Они не обладают пониманием контекста или общей картиной, что может привести к неправильным решениям. Алгоритмы могут давать сбой, если данные, на которых они обучаются, не являются представительными для широкого спектра ситуаций. Например, алгоритмы, получившие данные из определенной географической области, могут быть неэффективными при применении в других регионах. Кроме того, алгоритмы могут давать сбой в непредвиденных ситуациях. Например, если алгоритм для принятия решений в автомобиле полагается на данные с датчиков, и один из датчиков выходит из строя, это может привести к неправильным решениям. В-третьих, алгоритмы могут вызывать этические вопросы: усиливать неравенство, основанное на социальном классе, если алгоритм использует данные, которые отражают существующие предубеждения. В таблице 1 приведено краткое описание шагов, которыми стоит руководствоваться при применении алгоритмов.

Таблица 1 – План поэтапного использования алгоритмов принятия решений

№	Название шага	Ключевые сферы	Основные вопросы
---	---------------	----------------	------------------

1.	Согласие на использование алгоритмов конкуренции из-за	<ul style="list-style-type: none"> – Достижение намеченных целей; – Управление заинтересованными сторонами; – Понимание технологических возможностей; 	<ul style="list-style-type: none"> – Какие цели преследует организация и поможет ли алгоритм принятия решений достичь их? Если да, то как? – Кто является заинтересованными сторонами? Каковы их цели? Обладает ли организация необходимыми навыками
2.	Определение этичности данных	<ul style="list-style-type: none"> – Выявление ценностей организации; – Внедрение механизмов контроля, поддержки и обучения; – Планирование действий и их последствий; 	<ul style="list-style-type: none"> – Что значит «этично» для организации и клиентов? – Существует ли понимание последствий от использования алгоритмов? – Какое обучение необходимо для понимания, как работают алгоритмы? – Как можно обеспечить надлежащее управление данными и их использование?
3.	Управление осведомленностью клиентов	<ul style="list-style-type: none"> – Предоставление клиентам средств для контроля их данных; 	<ul style="list-style-type: none"> – Понимают ли клиенты, какие данные о них собираются и почему? – Знают ли клиенты законодательство, применяемое для алгоритмов?
4.	Гарантия использования корректных данных для алгоритмов	<ul style="list-style-type: none"> – Методы сбора данных для устранения неточностей и шумов; – Управление хранением и защитой данных; – Получение данных и разработка алгоритмов; 	<ul style="list-style-type: none"> – Какие данные необходимы для сбора? – Какие данные, которые собираются, не нужны? – Являются ли предоставляемые данные корректными и этичными?
5.	Оценка качества алгоритмов	<ul style="list-style-type: none"> – Оценка допустимости и прозрачности решений; – Оценка пригодности и предвзятости алгоритмов; – Обеспечение правдивости решений и преобладания здравого смысла; 	<ul style="list-style-type: none"> – Законно ли использовать алгоритмы принятия решений? – Подходят ли алгоритмы для решения задачи? – Является ли принятое решение логичным?

Алгоритмы принятия решений обладают большим потенциалом. Однако, если они используются без должного регулирования, это может привести к потере конфиденциальности и манипулированию.

Список использованных источников:

1. *Accountable Algorithms? The Ethical Implications of Data-driven Business Models // Journal of Service Management, 2020. – P. 163–185.*
2. *Ethical Issues in Big Data Analytics: A Stakeholder Perspective // Communications of the Association for Information Systems, 2019 – P. 718–747.*
3. *Smart Service Systems: An Interdisciplinary Perspective // Information Systems Journal, 2019 – P. 1201–1206.*