

УДК 004.051

**УСПЕШНЫЕ ПРИМЕРЫ ПРИМЕНЕНИЯ МАШИННОГО ОБУЧЕНИЯ
И ПРИЧИНЫ ИХ УСПЕХА****П.В. ДМИТРИЧЕНКО***(Представлено: канд. физ.-мат. наук, доц. Ю.Ф. ПАСТУХОВ)*

Изучены успешные проекты использующие машинное обучение на основании их анализа сделаны выводы о возможных технических причинах их успеха.

Введение. Деньги являются двигателем прогресса, а отсутствие возврата инвестиций эволюционным механизмом естественного отбора. Успешные проекты открывают дорогу тысячам подражателей и просто проектам которые могут заимствовать элементы успешного. Для понимания причин этого успеха необходим анализ технических условий в которых зарождались и росли наиболее успешные проекты использующие машинное обучение.

Финансовая часть является важной составляющей успеха данных проектов, но ввиду личного интереса автора внимание будет сконцентрировано на технических деталях, которые приводят к их успешному применению машинного обучения на практике.

Основная часть. Для начала кратко рассмотрим наиболее успешные проекты и стартапы использующие машинное обучение.

Финансовые проекты. Одними из наиболее перспективных проектов являются системы применяющие машинное обучение для вычисления счёта кредитования по которому определяются условия выдачи кредитов.

Affirm — предоставляет потребительский кредит для онлайн-покупок.

Upstart – платформа онлайн-кредитования, AI и machine learning применяет для оценки параметров кредита и автоматизации его получения. Проект работает и на b2b-рынке.

Здравоохранение является одной из областей активного развития машинного обучения.

Flatiron Health - оператор облачной платформы, специализированной на онкологии.

Сервисы, предоставляющие платформу для генерируемого пользователями контента или коммерческого. В данном контексте машинное обучение как правило используется для определения некачественного или сворованного контента, ранжирования, и генерации рекомендаций.

Генеральный директор YouTube, Сьюзен Войжитски, рассказала о том, что видео-сервис начнёт гораздо строже наказывать владельцев каналов, которые вредят его репутации. Так, будет реализовано введение штрафных санкций против авторов, которые «делают что-то вопиющее, что наносит значительный вред» всему сообществу. В частности, сервис наймёт для борьбы с нарушениями больше людей и начнёт ещё активнее использовать машинное обучение [1].

Youtube также использует машинное обучение для генерации рекомендаций и ранжирования этих рекомендаций.

Судя по недавнему патенту, Apple Watch смогут рекомендовать плейлисты из iTunes, которые идеально подойдут вашему сердечному ритму. Любая компания, которая работает с интернетом вещей, может использовать подобную технологию. Подключение нескольких устройств с одинаковым набором данных для обучения, сможет улучшить качество анализа, а значит и опыта клиента [4].

Проекты в сфере торговли являются одними из первых к которым было применено машинное обучение как правило в виде генерации рекомендаций, Amazon пошёл дальше этого и вводил персонализированные цены (данная схема также применена Ozon.ru и Dostavka.ru).

Amazon Personalize позволяет в режиме реального времени создавать индивидуально подобранные рекомендации по продуктам и контенту, а также целевые маркетинговые акции, чтобы активнее привлекать клиентов. В основе этого сервиса лежат 20 лет опыта работы компании Amazon в сфере рекомендаций и исследования с использованием машинного обучения. Сервис предоставляет в круглосуточное распоряжение клиентов систему рекомендаций на основе машинного обучения, подобной той, что используется на сайте Amazon.com [3].

Помимо персональных рекомендаций интересной фишкой на Amazon является персонализация цен. Стоимость товаров, которые похожи на те, что пользователь покупал ранее, значительно ниже стоимости других товаров из этой категории [4].

Крупнейшая в мире торговая площадка Alibaba широко использует машинное обучение и другие средства ИИ. В результате ее виртуальные витрины подстраиваются под каждого покупателя, а система поиска выдает для него наилучшие варианты. Чат-бот Ali Xiaomi может самостоятельно справиться с большинством обращений клиентов в техподдержку. [5].

Американская торговая сеть Target обнаружила, что с помощью машинного обучения можно предсказать не только поведение покупателей, но и перемены в их жизни, например, беременность. Алгоритмы Target работают настолько точно, что с помощью данных о покупках могут определить триместр беременности женщины, которая их совершала [4].

Далее будут рассмотрены возможные технические причины успешности вышеописанных и подобных им проектов.

Интеграция системы машинного обучения в уже существующие продукты, которые обладают большим объёмом собранных данных. Сбор данных используемых системами машинного обучения осуществляется уже существующими приложениями и сервисами. Консолидация из нескольких источников как это часто происходит в академических целях не происходит. Это избавляет такие проекты от множества проблем таких как неравномерность используемых данных, неравномерность детализации, а также различия в условиях, в которых данные собраны и используются.

Использование больших объёмов данных. Интеграция машинного обучения в уже существующие успешные и обладающие большими аудиториями продукты сильно упрощает сбор качественных данных если они уже не собраны и подготовку их для использования по методу больших данных. Кроме того, продукты обладающие большой аудиторией могут получать обратную связь в виде опросов пользователей об их удовлетворённости внедряемыми системами, а также получать об этом информацию исходя из поведения пользователей в приложении (например, вышел или продолжил пользоваться после показа контекстной рекламы, общее количество времени, проведённого в приложении после внедрения машинного обучения и т.п.).

Отсутствие рисков при получении неверных результатов. Большинство успешных проектов, использующих машинное обучение используют его для генерации рекомендаций или ранжирования существующего контента. Если подобные системы ошибаются владеющая ими компания не несёт больших финансовых потерь (в сравнении с оценкой экспертов или использования экспертных систем в том числе существующих в виде критериев на бумаге и/или формул) и как правило не может быть привлечена к легальной ответственности (исключения касаются законов о дискриминации и равенстве представления для, например, платформ, использующих машинное обучение для ранжирования контента). Это можно сравнить с экспертными системами, которые могут быть использованы для оценки состояния здоровья людей и выявления у них заболеваний, когда ложноотрицательная ошибка таких систем может привести к ухудшению здоровья пользователей вплоть до летальных исходов.

Вывод. Успешность внедрения машинного обучения в уже существующие информационные системы с большими объёмами данных собранных в этих системах, а не заимствованных из чужих наборов данных для решения задач не имеющих больших последствий при неверных единичных случаях обладает наибольшей перспективностью.

ЛИТЕРАТУРА

1. YouTube введёт санкции и машинное обучение для нейтрализации зловредных авторов [Электронный ресурс]. – Режим доступа: <https://hi-tech.ua/youtube-vvedyot-sanktsii-i-mashinnoe-obuchenie-dlya-neytralizatsii-zlovrednyih-avtorov/>. – Дата доступа: 15.09.2020.
2. Самые перспективные ИИ-стартапы [Электронный ресурс]. – Режим доступа: <https://vc.ru/finance/52042-samy-perspektivnye-ii-startapy>. – Дата доступа: 15.09.2020.
3. Amazon Personalize [Электронный ресурс]. – Режим доступа: <https://aws.amazon.com/ru/personalize/>. – Дата доступа: 15.09.2020.
4. Персонализация в Amazon.com, Ozon.ru и Dostavka.ru [Электронный ресурс]. – Режим доступа: <https://retailrocket.ru/blog/personalizatsiya-v-amazon-ozon-i-dostavka-ru/>. – Дата доступа: 15.09.2020.
4. 5 примеров использования машинного обучения в бизнесе [Электронный ресурс]. – Режим доступа: <https://rb.ru/story/machine-learning-in-business>. – Дата доступа: 15.09.2020.
5. Искусственный интеллект от Alibaba прошел экзамен в Стэнфорде лучше людей [Электронный ресурс]. – Режим доступа: <https://rb.ru/news/ai-alibaba-best/>. – Дата доступа: 15.09.2020.
6. 3 Criteria for a Successful Machine Learning Project [Электронный ресурс]. – Режим доступа: <https://indatalabs.com/blog/criteria-for-successful-machine-learning-project>. – Дата доступа: 15.09.2020.
7. Frame, Organize and Manage a Successful Machine Learning Project [Электронный ресурс]. – Режим доступа: <https://www.machinelearningmindset.com/organize-machine-learning-project/>. – Дата доступа: 15.09.2020.