

МАШИННОЕ ОБУЧЕНИЕ В ЦИФРОВОМ МАРКЕТИНГЕ

Жоль П. И.

Научный руководитель – ст. преподаватель Кондратёнок Е. В.

За последние 15 лет машинное обучение (machine learning, ML) получило широкое распространение. Многие из нас ежедневно используют приложения, в основе которых лежат технологии искусственного интеллекта (ИИ) и машинного обучения. Эти технологии уже стали причиной революции во многих отраслях [1].

Машинное обучение – это специализированный способ, позволяющий обучать компьютеры, не прибегая к программированию. Отчасти это похоже на процесс обучения младенца, который учится самостоятельно классифицировать объекты и события, определять взаимосвязи между ними.

ML открывает новые возможности для компьютеров в решении задач, ранее выполняемых человеком, и обучает компьютерную систему составлению точных прогнозов при вводе данных. Оно стимулирует рост потенциала искусственного интеллекта, являясь его незаменимым помощником, а в представлении многих даже синонимом.

Машинное обучение – одна из наиболее распространенных форм применения искусственного интеллекта современным бизнесом. Если компания еще не использует ML, то в ближайшее время наверняка оценит его потенциал, а ИИ станет основным двигателем IT-стратегии многих предприятий. Ведь искусственный интеллект уже сегодня играет огромную роль в трансформации развития IT-индустрии: клиенты больше внимания уделяют интеллектуальным приложениям, чтобы развивать свой бизнес с помощью ИИ. Он применим к любому рабочему процессу, реализованному в программном обеспечении, – не только в рамках традиционной деловой части предприятий, но также в исследованиях, производственных процессах и, во все большей степени, самих продуктах.

В основе машинного обучения лежат три одинаково важных компонента [1]:

1. Данные. Собираются всевозможными способами. Чем больше данных, тем эффективней машинное обучение и точнее будущий результат.
2. Признаки. Определяют, на каких параметрах строится машинное обучение.
3. Алгоритм. Выбор метода машинного обучения (при условии наличия хороших данных) будет влиять на точность, скорость работы и размер готовой модели.

В основу существования и развития машинного обучения легли три основных принципа [1]:

1. Инновационность: возможности ML открывают новые перспективы развития и роста практически всех отраслей экономики.
2. Специфичность: машинное обучение применяется для внедрения и разработки новых продуктов исключительно людьми, которые разбираются в IT-технологиях.

3. Простота: продукты, реализуемые с использованием технологий машинного обучения, становятся понятны даже школьникам и людям преклонного возраста.

Задачи, которые способно решить машинное обучение, напрямую определяют выгоды для бизнеса и возможности решения социальных проблем государствами разных стран.

В наше время искусственный интеллект до сих пор вызывает споры о безопасности для человека. Нужен ли ИИ маркетологам и какой и вред, и пользу приносит машинное обучение в современном маркетинге [2]?

С машинным обучением мы сталкиваемся чуть ли не каждый день. ИИ изучает наши интересы, наши предпочтения в музыке, кино, предлагает интересные нам новости, прокладывает быстрые и удобные для нас маршруты [2].

Проверка орфографии, сайты-знакомств, социальные сети, умные фильтры – это тоже машинное обучение. Для современных маркетологов это возможность выйти на новую аудиторию, показать нужный продукт нужным клиентам и продать его в интернете. ИИ собирает большие данные и дает возможность маркетологам быстро принимать важные решения.

Маркетологи собирают и сортируют огромные массивы данных. На сбор этих данных, их аналитику у компании уходят месяцы. Поэтому чаще всего машина сортирует и выбирает данные по определенным условиям, а человек строит теории и гипотезы на основе этих данных.

Можно выделить несколько направлений автоматизации маркетинга, которые актуальны прямо сейчас:

1. Анализ поведения клиента. Благодаря машинному обучению теперь можно узнать какую форму контента предпочитает пользователь, какой стиль и жанр музыки он слушает, на что подписан, какими книгами он интересуется, что хочет купить и т. д.

2. Автоматизированная визуализация данных. Любые схемы, диаграммы, графики, с помощью которых можно узнать данные о пользователе теперь стали намного удобнее и полезнее.

3. Последовательное планирование шагов. Как только система выполнила один блок задач, из него сразу же формируется другой. Работа идет непрерывно и постоянно.

Результаты применения машинного обучения видны в повседневной жизни:

1. Такси. Каждая поездка – это огромное количество данных: точка отправления, маршрут, время в пути, точка назначения, спрос. Анализ всего массива данных позволяет советовать пользователю такую точку отправления, до которой он сможет дойти пешком за несколько минут, водителю будет удобно подъехать, а сама поездка окажется намного дешевле [3].

2. Поисковые системы. Каждый раз, при обращении к поиску, начинают с работать множество алгоритмов. Одни исправляют в запросе ошибки или автоматически меняют язык раскладки, другие – ранжируют результаты поиска, третьи – анализируют запрос поиска.

3. Распознавание текста. Здесь машинное обучение может использоваться

в разных случаях. Во-первых, для определения шрифта. Существуют специальные программы, позволяющие распознавать почерк врача. Это нужно, чтобы все его записи автоматически попадали в систему, а не требовали кропотливого ручного ввода. При необходимости оцифровки бумажного документа, машина должна понимать шрифты.

4. Онлайн-переводчики. Об алгоритмах для переводов говорили много лет, но то возникали сложности с омонимией (слова одинаковые, а значения разные), то они плохо работали на предложениях разной длины. Последние результаты переводов на основе нейронных сетей понятны человеку и требуют лишь небольшой редактуры.

5. Обнаружение спама. В Gmail много лет идет сбор информации о видах писем пользователей (спам, не спам). Результатом этой работы стал алгоритм, обученный на этих данных, распознает спам и фишинговые атаки в 99,9 % случаев. Социальные сети тоже начинают перенимать опыт: Instagram. Разработаны алгоритмы для идентификации спама на русском, английском, испанском, португальском, арабском, французском, немецком, японском и китайском языках.

6. Голосовые помощники. На основе алгоритмов машинного обучения разработаны Siri, Google Assistant, Alexa, Алиса.

7. Банки и платежные системы. Для банковских алгоритмов каждый клиент банка является набором характеристик. Анализ характеристик позволяет принимать решения о возможности выдачи кредита. Сравнивая миллионы транзакций между собой, алгоритмы PayPal ищут подозрительные обращения и блокируют их.

8. Рекомендательные системы. Мощные алгоритмы разрабатываются в Netflix. На основе анализа просмотров пользователя алгоритмы предлагают подходящий по предпочтениям фильм или сериал [3].

Маркетологи используют машинное обучение для нахождения паттернов в действиях пользователей на сайте. Это помогает им предугадывать дальнейшее поведение других пользователей и быстро изменять рекламные предложения [4].

Паттерн – это набор поведенческих реакций или последовательностей на действие. Можно рассмотреть пример паттерна на сайте, если пользователя не заинтересовало предложение в модальном окне, у него есть несколько вариантов как его закрыть.

1. Кликнуть на крестик.
2. Кликнуть на рабочую область мимо модального окна.
3. Ввести данные и пойти дальше.

Так же есть еще параметр времени от всплытия окна до его закрытия. Таким образом появляется четыре параметра данных о пользователе.

1. Клик на крестик – может принимать значения true/false.
2. Клик мимо модального окна – может принимать значения true/false.
3. Клик на поле ввода данных – может принимать значения true/false и собирать те данные, которые ввел пользователь.
4. Время просмотра окна – 5 секунд.

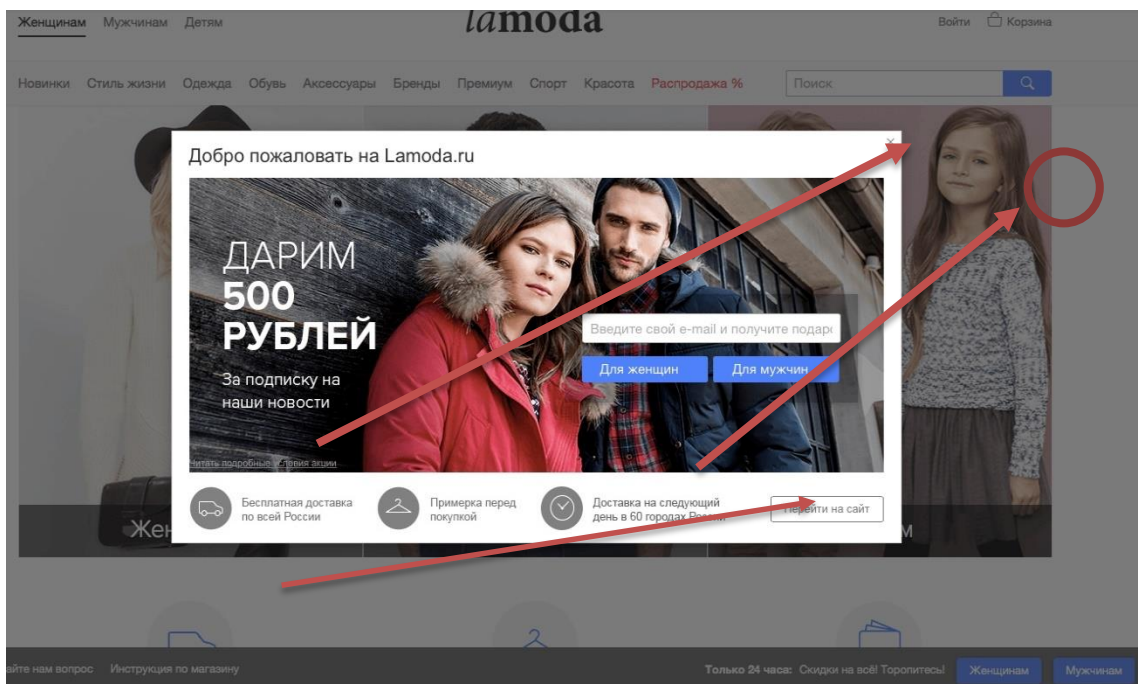


Рис. 1 – Модальное окно

Когда собирается много таких параметров, данные обретают целостность, т. к. они содержат паттерны поведения и связь между ними. Это дает возможность дополнить данные о пользователе недостающими параметрами, исходя из тех данных, которые у нас уже есть по другим пользователям.

Основные преимущества использования машинного обучения в маркетинге [3]:

- повышение качества анализа данных;
- возможность анализа большого объема данных за меньшее количество времени;
- адаптируемость под изменения и ввод новых данных;
- автоматизация маркетинговых процессов;

Примеры использования машинного обучения в маркетинге:

1. Система рекомендаций.

Рекомендательная система – система, которая может предложить пользователю товар, в котором он наиболее заинтересован в текущее время с целью повышения вероятности совершения покупки. Обычно эти данные используются в блоках «Рекомендуемые товары» и «Похожие товары», в email и push – рассылках, в социальных сетях. В результате пользователи видят специальные персонализированные блоки, которые им интересны [3].

Пример рекомендательного блока:

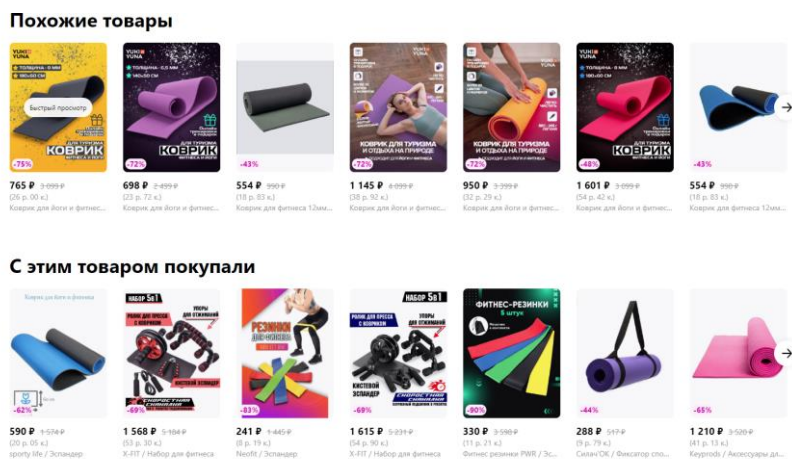


Рис. 2 – Блок «похожие товары»

2. Прогнозируемый таргетинг.

Компания расходует свой бюджет на пользователей своей целевой аудитории и не расходует на тех, кто не попадает в эту целевую аудиторию.

Есть два вида наиболее используемого таргетинга:

- таргетинг по сегментам – показ рекламы группам пользователей с похожими наборами атрибутов и описаний;
- триггерный таргетинг – показ рекламы пользователям после совершения ими определенного действия (просмотр продукта, просмотр товара, добавление товара в избранное, добавление товара в корзину, покупка товара).

В прогнозируемом таргетинге используются все возможные комбинации десятков или сотен параметров пользователя со всеми возможными значениями, а в остальных видах таргетинга используется ограниченное количество параметров с определенными диапазонами значений.

Пример: создаются сегменты на основе вероятности совершения покупки. Сегменты загружаются в Google Ads, Facebook Ads, Яндекс. Аудитории и другие рекламные системы, а потом используются для запуска рекламных кампаний.

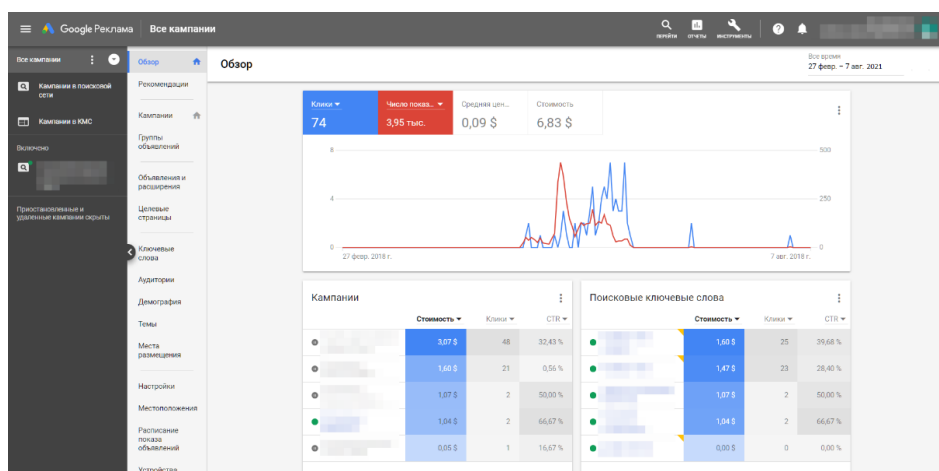


Рис. 3 – «Рекламная компания на Google ADS»

3. Копирайтинг и контент.

Машинный интеллект анализирует данные и контент сайта: кол-во слов в блоке на лендинге, обороты слова, которые влияют на клики и т. д. Некоторые зарубежные компании предоставляют сервисы, где машина сама пишет текст (когнитивный контент). Для этого она использует математику, статистику, лингвистику и данные об эмоциях. Система изучает слова, которые вызывают отклик клиентов, и использует их для вовлечения или конверсии.

Применять машинное обучение к бизнесу без понимания того, что действительно работает, а что нет, нецелесообразно. Однако использовать алгоритмы машинного обучения важно уже сейчас. Эксперты прогнозируют, что технологии машинного обучения и решения для сбора, хранения и «умного» анализа данных могут дать бизнесу заметный толчок и конкурентные преимущества в борьбе за пользователя [4].

4. Системы навигации.

Данные о местоположении – это важная категория данных, с которыми часто приходится иметь дело в проектах машинного обучения. Они, как правило, дают дополнительный контекст к данным используемого приложения [5].

Имея точные координаты достопримечательностей, можно рассчитать расстояние между ними. Например, важной информацией может быть расстояние до ближайшей больницы или расстояние между местами посадки и высадки в каршеринге. Такое расстояние можно измерить несколькими способами:

1. Эвклидово расстояние – простое расстояние по прямой между координатами (x, y) двух точек. Это расстояние измеряется на плоской 2D-поверхности.

2. Геодезическое расстояние – измеряется на сферической Земле, то есть на трехмерной поверхности. Например, кратчайшим расстоянием будет расстояние между двумя точками на сфере. Расстояние Haversine – это примерно то же, что и дуга большого круга, но для его расчета используется формула Haversine.

3. Манхэттенское расстояние – применяется к городским местоположениям, в которых улицы образуют кварталы. Используется для расчета фактического расстояния (например, при вождении автомобиля или пешей прогулке) между двумя точками вдоль городских улиц. Такой расчет более логичен, чем расчет расстояния по прямой. Название расстояния происходит от района Манхэттен в Нью-Йорке, заполненного кварталами квадратной формы с дорогами, идущими параллельно друг другу и пересекающимися под прямым углом. Однако на практике, хотя улицы и пролегают прямо, их направление не всегда в точности ориентировано на север, юг, восток или запад. Это обстоятельство учитывается: рассчитывается скорректированное расстояние с учетом угла поворота карты улиц города.

Геолокационные данные и связанные с ними характеристики могут быть представлены в виде таблиц с числовыми или дискретными переменными. Например, если были получены данные о местоположении дома, то наравне со всеми другими атрибутами, не связанными с местоположением (например, временем постройки и площадью дома), могут использоваться геодезические данные.

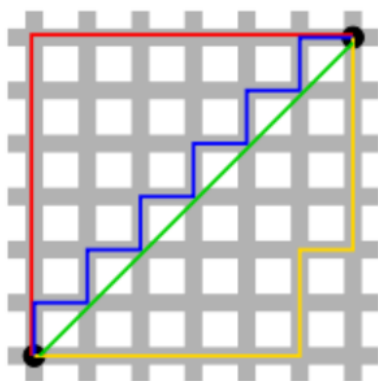


Рис. 4 – Манхэттенское расстояние

Другими словами, работа с данными о местоположении в основном связана с подготовкой данных, а не с созданием каких-либо специфических для данной местности моделей машинного или глубокого обучения.

Машинное обучение не только существенно облегчает работу маркетологов, выполняя рутинные операции. Оно помогает систематизировать, обрабатывать и анализировать большие объемы данных. Машинное обучение позволяет моделировать различные ситуации, делать прогнозы для эффективных маркетинговых решений, с его помощью можно скорректировать стратегию развития бизнеса.

Литература

1. Машинное обучение: просто о сложном [Электронный ресурс]: информационный ресурс SberCloud. URL: <https://sbercloud.ru/ru/warp/machine-learning-about> (дата обращения: 08.04.2022).

2. Вторжение машинного обучения: от заказа такси до персонализации контента [Электронный ресурс]: информационный ресурс блог.ingate. URL: <https://blog.ingate.ru/detail/vtorzhenie-mashinnogo-obucheniya-ot-zakaza-taksi-do-personalizatsii-kontenta/> (дата обращения: 08.04.2022).

3. Домингос, П. Верховный алгоритм. Как машинное обучение изменит наш мир / П. Домингос. – М.: Манн, Иванов и Фербер, 2016. – 56 с.

4. Использование геолокационных данных в машинном обучении: основные методы [Электронный ресурс]: информационный ресурс HABR.COM. URL: <https://habr.com/ru/company/skillfactory/blog/554438/> (дата обращения: 08.04.2022).