

УДК 338.23:004 + 339.13:004.438.5

О ПОВЫШЕНИИ КОНВЕРСИИ ПРОДАЖ ИНТЕРНЕТ-РЕСУРСОВ ПУТЕМ РАСЧЕТА ВЕРОЯТНОСТИ СОВЕРШЕНИЯ КЛЮЧЕВОГО ДЕЙСТВИЯ КЛИЕНТОМ

Р.А. Мусеев, А.В. Мокеев, В.В. Мокеев

В работе рассматривается задача создания новых и масштабирования аналитических Web-сервисов для повышения конверсии ключевых показателей бизнеса с большим количеством интернет-клиентов за счет разработки и применения новых методов машинного обучения для прогнозирования приоритетов коммуникаций с клиентами. В настоящее время разработан прототип аналитического Web-сервиса, который внедрен в работу одной из крупнейших организаций, занимающейся бизнес-образованием в России.

Ключевые слова: клиент, продажа, вероятность покупки, бустинг.

В настоящее время все большее количество данных о жизни и активности моментально оцифровывается, появляется принципиально иная возможность автоматизированного интеллектуального взаимодействия с людьми и обществом на основе анализа их активности. Современные методики анализа данных позволяют выявлять эффективные комбинации действий для достижения личностных, общественных и коммерческих задач. С начала 2000-х сформировалось направление «Большие данные», так как стало понятно, что важно не количество данных, а возможность извлечь из этих данных ценный опыт. Поэтому основное современное направление анализа «Машинное обучение» в настоящее время ассоциируется с направлением «Большие данные» как основным практически применимым направлением работы с данными. Одним из важных источников эффективности компании являются ее клиенты, так как большая часть денежного потока формируется именно из платежей клиентов. В связи с этим задача создания системы эффективной работы с клиентами становится приоритетной для современного бизнеса. На растущем рынке предприятие ставит в качестве приоритетной задачи привлечение новых клиентов. Менеджеры в первую очередь интересуются объемом и динамикой продаж, завоевание доли рынка. Компания тратит средства на привлечение клиента, маркетинговые бюджеты максимальны. При этом очень часто компании готовы взаимодействовать с любым, даже не очень прибыльным клиентом. Количество привлеченных клиентов и занимаемая доля рынка являются главными индикаторами результативности маркетинговых процессов на стадии привлечения клиентов. Однако по мере насыщения рынка и дости-

жения уровня самокупаемости приоритетом становится задача не столько привлечения новых клиентов, сколько удержания существующих. Главным индикатором результативности становится уровень удержания клиентов. Высокий уровень удержания, как правило, свидетельствует о том, что клиенты лояльны к компании и предпочитают оставаться с компанией в будущем. В условиях высокой конкуренции важно использовать каждый контакт с клиентом. Воронка продаж – это один из инструментов работы с клиентами, позволяющий оценивать результаты на каждой стадии продажи. Она показывает, сколько клиентов, поступающих на вход, попадают в дальнейшую обработку и в конечном счете доходят до этапа заключения сделки. Иными словами, на каждом этапе продаж, в один период времени менеджер работает с определенным количеством клиентов, по завершении которого часть клиентов «отсеивается», а часть переходит на следующий этап, где ситуация повторяется до тех пор, пока не состоится продажа. На стадии привлечения задача заключается в формировании входа воронки таким образом, чтобы на каждом последующем этапе отсев клиентов был минимален (повышение коэффициента конверсии). Воронка продаж должна формироваться как многоступенчатый процесс, при котором на каждом этапе до момента продажи определяются коэффициенты конверсии, показывающие полезный результат (эффект) на единицу затраченных усилий по привлечению. К таким коэффициентам относятся: коэффициент конверсии обращений, коэффициент конверсии контактов, коэффициент конверсии коммерческих предложений, коэффициент конверсии заказов. В долгосрочной перспективе компания стремится максимизировать период сотрудничества с клиентом. Для большинства руководителей вполне очевидно, что приверженность покупателей имеет решающее значение для делового успеха. Эту стадию можно назвать стадией развития взаимоотношений с клиентом. Создание устойчивых партнерских отношений с лояльными клиентами приносит большую прибыль, чем единичные транзакции и максимизация прибыли по каждому отдельному контракту. Если работа с клиентами учитывает историю взаимоотношений именно с этим клиентом, а также опирается на модель поведения других клиентов, относящихся к тому же сегменту, то вероятность успеха, несомненно, возрастает.

В связи с этим задача создания системы эффективной работы с клиентами становится приоритетной для современного бизнеса. Именно ради нее и реализуется сервис, который направлен в первую очередь на решение задач определения ценности клиента и потенциальных клиентов. Для этого решается задача определения вероятности совершения ключевого действия клиентом с использованием алгоритмов градиентного и экстремального бустинга. Технология решения задач включает четыре этапа. На первом этапе формируются данные, которые представляют собой многомерный лог, фиксирующий активность пользователя на сайте и его действия, свя-

занные с оформлением и оплатой договоров. Для каждого пользователя имеется несколько записей. Ввиду количественной избыточности и практической сложности интерпретации данных используется информация о тех пользователях, для которых имеется возможность учесть динамику в их поведении. На втором этапе создается модель для решения задачи с использованием бустинговых методов.

С помощью математических моделей выявляются закономерности в поведении клиентов и предсказывается, как они поведут себя в будущем. Основное внимание уделяется модели классификации. Модель классификации описывает правила или набор правил, в соответствии с которыми можно отнести описание любого нового объекта к одному из классов. Такие правила строятся на основании информации о существующих объектах путем разбиения их на классы. Модели разрабатываются в первую очередь для решения задач определения вероятностей первой и последующих покупок. В рамках задачи определения вероятностей первой покупки модель оценивает данные клиентов, такие как: клики на письма *email*-рассылки и взаимодействие с сайтом, чтобы определить, кто из них и насколько готов совершить первую покупку. Имея такую информацию о клиентах, можно изменить систему скидок и сформировать выгодные предложения для увеличения прибыли. Если вероятность того, что клиент готов совершить покупку, высокая, то клиент не нуждается в предложении больших скидок. В случае, если клиенты сомневаются – покупать или нет (вероятность не высокая), то это может означать, что подтолкнуть его к покупке могут выгодные предложения. Если клиент сделал первую покупку, то целью становится построение модели оценки вероятности следующих покупок. Такая модель анализирует первые покупки, чтобы определить вероятность покупок клиентов в будущем. Это позволяет определить клиентов, которых нужно подтолкнуть к покупкам путем поощрений (подарки или скидки).

На сегодняшний день бустинговые алгоритмы являются одними из самых мощных алгоритмов классификации.

Это достигается благодаря адаптивной технике построения композиции. К тому же бустинг предоставляет множество возможностей для вариаций. Во-первых, можно рассматривать различные функции потерь. Это позволяет решать как задачи классификации, так и задачи регрессии. А возможность выбора произвольной функции потерь позволяет акцентировать внимание на особенностях анализируемых данных. Во-вторых, возможно рассмотрение любого семейства базовых алгоритмов. А это, опять же, дает широкие возможности учета особенностей данной задачи. Бустинг над решающими деревьями считается одним из наиболее эффективных вариантов бустинга. А учитывая, что решающие деревья в свою очередь тоже используют базовые алгоритмы (например, пороговые, линейные и т.п.), в результате получается огромное количество вариантов для настрой-

ки. В-третьих, благодаря достаточной простоте метода и четкому математическому обоснованию в каждой конкретной вариации бустинга несложно провести некоторые математические и алгоритмические оптимизации, которые заметно ускорят работу алгоритма. В случае, если на вход модели подаются разреженные данные (с большим количеством нулей и пустых ячеек), то бустинговые алгоритмы используют оптимизированные структуры для хранения разреженных данных, что позволяет существенно экономить память ЭВМ. На третьем этапе оценивается качество модели. Для задачи распознавания с неуравновешенными классами используются такие метрики, как точность и полнота. ROC-кривая представляет зависимость, позволяющую оценить качество бинарной классификации. ROC-кривая отображает соотношение между долей объектов от общего количества носителей признака, верно классифицированных (*true positive rate*, TPR), и долей объектов от общего количества объектов, ошибочно классифицированных (*false positive rate*, FPR). Количественную интерпретацию ROC даёт показатель AUC – площадь, ограниченная ROC-кривой и осью доли ложных положительных классификаций. Чем выше показатель AUC, тем качественнее классификатор, при этом значение 0,5 демонстрирует непригодность выбранного метода классификации (соответствует случайному гаданию). Для того чтобы избежать переобучения модели, добавляется функция автоматической кросс-валидации. В результате разработанный сервис предоставляет интерфейс для предсказания вероятности продаж на потоке данных заказчика. На четвертом этапе разработанная модель используется для решения задач. Предприняты реальные шаги, начиная с 2015 года, по продвижению сервиса в работу образовательной организации. Использование данного сервиса позволило повысить конверсии продаж образовательной организации. Результаты подтверждены улучшением показателей компании, подписанным договором на использование сервиса. Результаты использования показывают, что происходит увеличение потока сделок и оборота на 5–10 % и выше. Компании-заказчики готовы платить от 5 до 20 % от объема собственной прибыли, которая была получена от клиентов, выявленных с использованием сервиса.

[К содержанию](#)