

УДК 004.08

DOI: [10.26102/2310-6018/2026.57.6.018](https://doi.org/10.26102/2310-6018/2026.57.6.018)

Применение uplift-моделирования для повышения эффективности адресных маркетинговых кампаний в ритейле

Т.В. Азарнова✉, Ю.И. Реброва

Воронежский государственный университет, Воронеж, Российская Федерация

Резюме. Рост конкуренции на рынке розничной торговли и электронной коммерции требует от компаний более точного подхода к планированию маркетинговых кампаний и персонализации коммуникаций с клиентами. Существующие модели предсказания отклика не способны изолировать эффект маркетингового воздействия от естественного покупательского поведения, что ведет к нерациональному расходованию бюджетов и затрудняет объективную оценку проводимых кампаний. Решением данной проблемы выступает uplift-моделирование – подход, позволяющий оценивать причинно-следственный эффект воздействия на уровне отдельного потребителя и определять сегменты аудитории с наибольшей восприимчивостью к коммуникациям. Настоящая статья посвящена сравнительному анализу методов uplift-моделирования с целью выбора наиболее результативного из них для оценки эффекта маркетингового воздействия. В рамках исследования рассмотрены пять методов (S-Learner, T-Learner, Class Transformation, X-Learner и R-Learner) на открытых данных Lenta Uplift Modeling Dataset, предоставленных розничной сетью Лента в рамках хакатона BigTarget совместно с Microsoft. Качество моделей оценивалось при помощи специализированных метрик (Uplift@k, Qini AUC, Uplift AUC, Weighted Average Uplift, Average Squared Deviation). По результатам анализа выявлены достоинства и недостатки каждого из подходов, а также установлен метод, показавший наилучшие результаты на рассматриваемом наборе данных.

Ключевые слова: uplift-моделирование, машинное обучение, оценка эффекта воздействия, таргетированный маркетинг, персонализация коммуникаций, метрики качества uplift-моделей.

Для цитирования: Азарнова Т.В., Реброва Ю.И. Применение uplift-моделирования для повышения эффективности адресных маркетинговых компаний в ритейле. *Моделирование, оптимизация и информационные технологии*. 2026;14(6). URL: <https://moitvvt.ru/ru/journal/article?id=2345> DOI: 10.26102/2310-6018/2026.57.6.018

Application of uplift modeling to enhance the effectiveness of targeted marketing campaigns in retail

T.V. Azarnova✉, J.I. Rebrova

Voronezh State University, Voronezh, the Russian Federation

Abstract. The growing competition in the retail and e-commerce markets requires companies to adopt more precise approaches to planning marketing campaigns and personalizing customer communications. Existing response prediction models fail to isolate the effect of marketing interventions from natural purchasing behavior, leading to irrational budget spending and complicating the objective evaluation of campaigns. Uplift modeling emerges as a solution – an approach that enables causal effect assessment at the individual consumer level and identifies audience segments most responsive to communications. This article presents a comparative analysis of uplift modeling methods to select the most effective one for evaluating marketing impact. The study examines five methods (S-Learner, T-Learner, Class Transformation, X-Learner, and R-Learner) using the open Lenta Uplift Modeling Dataset provided by the Lenta retail chain during the BigTarget hackathon in collaboration with Microsoft. Model performance was evaluated using specialized metrics (Uplift@k, Qini AUC, Uplift AUC, Weighted

Average Uplift, Average Squared Deviation). The analysis reveals the strengths and weaknesses of each approach and identifies the top-performing method for this dataset.

Keywords: uplift modeling, machine learning, treatment effect evaluation, targeted marketing, communication personalization, uplift model quality metrics.

For citation: Azarnova T.V., Rebrova J.I. Application of uplift modeling to enhance the effectiveness of targeted marketing campaigns in retail. *Modeling, Optimization and Information Technology*. 2026;14(6). (In Russ.). 2026;14(6). URL: <https://moitvvt.ru/ru/journal/article?id=2345> DOI: 10.26102/2310-6018/2026.57.6.018

Введение

Усиление конкурентного давления в ритейле и активный рост онлайн-торговли обусловили повышенный интерес к персонализированным маркетинговым кампаниям как инструменту управления клиентской базой. Тем не менее практическая эффективность подобных кампаний остается ограниченной, поскольку показатели отклика традиционно не превышают нескольких процентов, а рентабельность инвестиций удерживается на умеренном уровне. Ключевой причиной служит принципиальное ограничение классических методов сегментации, направленных на предсказание общей вероятности целевого действия, но неспособных отделить органическое поведение клиента от эффекта маркетингового воздействия. Следствием этого становится систематическое нерациональное расходование бюджета на клиентов, совершающих покупку независимо от коммуникации, и на тех, кто остается к ней нечувствителен.

Эволюция маркетинговой аналитики в ритейле прошла путь от предсказательных моделей к причинно-следственному анализу. Look-alike модель [1] позволяет находить клиентов, схожих с целевой аудиторией по демографическим и поведенческим характеристикам, и эффективна для привлечения новой аудитории, однако не учитывает влияние самой коммуникации на поведение клиента и не разграничивает естественную склонность к покупке от эффекта воздействия. Response-модель на базе логистической регрессии обеспечивает более высокую прогностическую точность за счет обучения на эмпирических данных реальных кампаний и хорошо справляется с задачей массового таргетинга, однако принципиально не решает проблему разграничения органического поведения клиента и эффекта коммуникации [2, 3], что снижает экономическую эффективность кампаний и приводит к неоправданным затратам бюджета.

Осознание данных ограничений привело к появлению uplift-моделирования как самостоятельного направления в маркетинговой аналитике, ориентированного на измерение инкрементального эффекта воздействия на уровне отдельного клиента. Основы направления заложены в работе¹, предложившей метод Significance-Based Uplift Trees с разделением клиентов на четыре поведенческих типа: «убеждаемые», «независимые покупатели», «безразличные» и «негативно реагирующие». Данный подход был расширен в работе [4], в которой предложены деревья решений для uplift-моделирования с поддержкой единственного и множественного воздействий. Практическая ценность подхода состоит в концентрации бюджета кампании исключительно на убеждаемых клиентах, что позволяет существенно повысить отдачу от коммуникаций. Дальнейшее развитие направления связано с классом методов meta-learners: в работе [5] систематизированы архитектуры S-Learner, T-Learner и X-Learner, различающиеся способом оценки каузального эффекта на уровне отдельного клиента и

¹ Radcliffe N.J., Surry P.D. *Real-World Uplift Modelling with Significance-Based Uplift Trees*. Portrait Technical Report TR-2011-1. Stochastic Solutions. URL: <https://stochasticsolutions.com/pdf/sig-based-up-trees.pdf> (дата обращения: 12.02.2026).

степенью устойчивости к дисбалансу выборок. Сравнительный анализ методов uplift-моделирования на синтетических и реальных данных [6] показал, что совместное применение нескольких подходов к оценке каузального эффекта повышает точность сегментации аудитории по восприимчивости к воздействию. Исследования применения uplift-моделирования в задачах прямого маркетинга [7] подтверждают его превосходство над response-моделями по показателям прироста прибыли и демонстрируют, что фокус на инкрементальном эффекте позволяет существенно повысить отдачу от коммуникаций.

Несмотря на растущий интерес к uplift-моделированию, систематическая оценка существующих подходов на реальных розничных данных с применением комплекса специализированных метрик остается недостаточно изученной областью. Целью настоящей работы является сравнительный анализ ведущих uplift-методов применительно к задачам оптимизации маркетинговых коммуникаций в ритейле. Для достижения данной цели в работе рассматриваются и сопоставляются подходы S-Learner, T-Learner, Class Transformation, X-Learner и R-Learner. Оценка качества каждой модели проводится на основе комплекса специализированных метрик: Uplift@k, Qini AUC, Uplift AUC, Weighted Average Uplift и Average Squared Deviation, обеспечивающих всестороннее измерение как ранжирующей способности модели, так и точности абсолютных предсказаний инкрементального эффекта. Информационной базой данного исследования является открытый набор данных Lenta Uplift Modeling Dataset, предоставленный розничной сетью «Лента».

Материалы и методы

Под uplift-моделированием понимается метод машинного обучения, задачей которого является оценка ожидаемого прироста в целевом поведении клиента вследствие маркетингового воздействия. Использование этого инструмента дает возможность эффективнее распределять бюджет рекламной кампании, направляя средства на ту аудиторию, чья активность напрямую обусловлена коммуникацией, одновременно исключая из охвата нецелевые сегменты [7].

Методология uplift-моделирования основана на оценке количественной разницы в целевом поведении клиентов в зависимости от того, было ли к ним применено маркетинговое воздействие. Для формального описания этого эффекта вводится величина τ_i , называемая causal effect – индивидуальный причинно-следственный эффект воздействия, отражающий, насколько именно маркетинговая коммуникация изменила вероятность целевого действия для конкретного клиента. Данный показатель вычисляется через разность вероятностей наступления целевого события в зависимости от факта воздействия:

$$\tau_i = Y_i^1 - Y_i^0, \quad (1)$$

где Y_i^1 – вероятность отклика при наличии коммуникации, Y_i^0 – вероятность отклика при ее отсутствии.

Принципиальная сложность в оценки τ_i состоит в том, что невозможно одновременно наблюдать оба исхода для одного клиента, так как он либо получает коммуникацию, либо нет. В прикладных исследованиях используют его условное математическим ожиданием, называемым Conditional Average Treatment Effect (CATE):

$$CATE = E[Y_i^1|X_i] - E[Y_i^0|X_i], \quad (2)$$

где X_i – вектор признаков клиента i .

Поскольку ни τ_i , ни $CATE$ для i -го клиента непосредственно наблюдать невозможно, в качестве практической оценки используется эмпирическая оценка $CATE$, выражаемая через наблюдаемые исходы следующим образом:

$$uplift = \widehat{CATE} = E[Y_i | X_i = x, W_i = 1] - E[Y_i | X_i = x, W_i = 0], \quad (3)$$

где Y_i – наблюдаемый отклик клиента, определяемый его групповой принадлежностью:

$$Y_i = W_i Y_i^1 + (1 - W_i) Y_i^0 = \begin{cases} Y_i^1, & \text{если } W_i = 1 \\ Y_i^0, & \text{если } W_i = 0 \end{cases} \quad (4)$$

где $W_i = 1$ означает, что клиент входит в группу воздействия (treatment), $W_i = 0$ – клиент относится к контрольной группе (control).

Корректность применения данной формулы опирается на выполнение предположения об условной независимости (Conditional Independence Assumption – CIA), в соответствие с которым потенциальные исходы клиента $\{Y_i^0, Y_i^1\}$ не зависят от факта его принадлежности к группе воздействия или контрольной группе при условии наблюдаемых признаков X_i :

$$CIA : \{Y_i^0, Y_i^1\} \perp\!\!\!\perp W_i | X_i. \quad (5)$$

Иными словами, распределение клиентов между группами должно быть случайным и не определяться значениями их признаков. При соблюдении данного условия наблюдаемая реакция Y_i зависит от принадлежности к группе лишь через факт воздействия, что обеспечивает корректную идентификацию причинно-следственного эффекта из наблюдаемых данных.

Реализация uplift-моделирования на практике предполагает выбор конкретного алгоритма оценки uplift. В контексте ритейла, где клиентские базы отличаются высокой неоднородностью, а данные рандомизированных экспериментов нередко содержат существенный дисбаланс между группами воздействия и контроля, выбор подходящего метода оказывает значительное влияние на качество сегментации аудитории и, как следствие, на эффективность распределения маркетингового бюджета. Для оценки uplift существует ряд подходов, каждый из которых может быть адаптирован к конкретным условиям эксперимента в зависимости от размера и баланса групп, типа целевой переменной и особенностей клиентских данных.

Single-model approach или S-learner является наиболее простым подходом в классе meta-learners. Его ключевая идея состоит в том, что задача оценки инкрементального эффекта не выделяется в отдельную подзадачу, а решается в рамках единой модели. Индикатор принадлежности к группе воздействия W_i включается в качестве обычного признака наравне с остальными характеристиками клиента, после чего на всей объединенной выборке обучается единая модель $\hat{\mu}(x, w)$. Оценка uplift для конкретного клиента получается как разность предсказаний этой модели при $W_i = 1$ и $W_i = 0$.

$$uplift(x) = \hat{\mu}(x, 1) - \hat{\mu}(x, 0). \quad (6)$$

Подход отличается простотой реализации и совместимостью с любым базовым алгоритмом машинного обучения без какой-либо модификации его архитектуры. Вместе с тем алгоритмы типа случайного леса или градиентного бустинга могут полностью игнорировать индикатор воздействия W_i при построении разбиений, если он оказывается менее информативным, чем остальные признаки клиента, что приводит к смещению оценок uplift в сторону нуля и снижает способность модели корректно разграничивать клиентов из treatment и control групп.

Two model approach или T-Learner устраняет указанное ограничение путем явного разделения обучающей выборки на две независимые части и обучения двух отдельных моделей. Первая модель $\hat{\mu}_1(x)$ обучается на данных группы воздействия, вторая $\hat{\mu}_0(x)$ – на данных контрольной группы. Каждая решает обособленную задачу: оценить вероятность целевого события в своих условиях, не смешивая сведения о клиентах с коммуникацией и без нее. Итоговая оценка uplift для каждого клиента вычисляется как разность предсказаний двух моделей:

$$uplift(x) = \hat{\mu}_1(x) - \hat{\mu}_0(x). \quad (7)$$

Обучение каждой модели исключительно на данных своей группы позволяет корректно изолировать влияние клиентских признаков от эффекта воздействия, что повышает точность оценки индивидуальной реакции на коммуникацию. Тем не менее при существенном дисбалансе размеров групп данный подход также имеет свои ограничения. Поскольку итоговая оценка uplift формируется как разность предсказаний двух независимых моделей, их ошибки суммируются, а не компенсируют друг друга.

Метод Class Transformation, предложенный [8], представляет собой математически обоснованный подход, принципиально отличающийся от двух предыдущих тем, что он не требует обучения отдельных моделей для каждой группы и не полагается на косвенное сравнение их предсказаний. Вместо этого предлагается преобразовать исходную задачу путем введения новой целевой переменной Z_i , которая конструируется таким образом, чтобы ее условное математическое ожидание непосредственно соответствовало искомому эффекту воздействия. В базовой постановке новая переменная определяется следующим образом:

$$Z_i = Y_i \cdot W_i + (1 - Y_i) \cdot (1 - W_i), \quad (8)$$

где Z_i представляет собой трансформированную целевую переменную для i -го клиента, Y_i отражает наблюдаемый отклик i -го клиента на маркетинговое воздействие, W_i является бинарным индикатором принадлежности к группе: значение $W_i = 1$ соответствует клиентам из treatment группы, $W_i = 0$ соответствует клиентам из control группы.

При условии равного разбиения выборки на treatment и control группы, то есть

$$P(W = 1) = P(W = 0) = \frac{1}{2}, \quad (9)$$

условное математическое ожидание трансформированной переменной Z_i оказывается непосредственно связано с искомым эффектом воздействия. Следовательно, оценка uplift выражается через предсказание новой целевой переменной следующим образом:

$$uplift = 2 \cdot P(Z = 1) - 1. \quad (10)$$

X-Learner представляет собой метод с перекрестным обучением двух моделей, разработанный для ситуаций с существенным дисбалансом между treatment и control группами, при котором стандартные подходы утрачивают точность вследствие непропорционального распределения наблюдений [5]. Ключевое отличие данного метода от T-Learner заключается в том, что модели не изолированы друг от друга, а каждая из них задействует предсказания противоположной группы.

Построение X-Learner предполагает четыре последовательных этапа. На первом параллельно обучаются две независимые модели отклика: $\hat{\mu}_0(x)$ на данных контрольной группы и $\hat{\mu}_1(x)$ на данных группы воздействия. На втором этапе целевые переменные каждой из групп преобразуются с опорой на предсказания модели, обученной на данных

другой группы. В результате для клиентов группы воздействия вменяемый эффект принимает значение \tilde{D}_i^1 , а для клиентов из контрольной группы – \tilde{D}_i^0 :

$$\tilde{D}_i^1 = Y_i^1 - \hat{\mu}_0(X_i^1), \quad (11)$$

$$\tilde{D}_i^0 = \hat{\mu}_1(X_i^0) - Y_i^0. \quad (12)$$

Если бы оценки $\hat{\mu}_0$ и $\hat{\mu}_1$ были реальными, а не предсказанными величинами, то \tilde{D}_i^1 и \tilde{D}_i^0 в точности совпадали бы с истинным uplift каждого клиента. На третьем этапе на преобразованных таргетах \tilde{D}_i^1 и \tilde{D}_i^0 обучаются две новые модели $\hat{\tau}_1(x)$ и $\hat{\tau}_0(x)$ соответственно. На последнем шаге итоговая оценка uplift формируется как взвешенная сумма предсказаний этих двух моделей [4]:

$$uplift(x) = g(x) \cdot \hat{\tau}_0(x) + (1 - g(x)) \cdot \hat{\tau}_1(x), \quad (13)$$

где $g(x) \in [0; 1]$ является весовой функцией, в качестве которой рекомендуется использовать оценку propensity score. При большом размере treatment группы относительно control рекомендуется выбирать $g(x)$ близким к 1, при обратном соотношении – близким к 0. Примечательно, что $g(x)$ не обязана быть константной. Ее вполне можно задать как функцию признаков отдельного клиента. В таком случае модель получает дополнительный запас гибкости, что особенно ценно в случае неоднородной клиентской базы ритейлера.

R-Learner опирается на преобразовании Робинсона и принцип остаточной регрессии [9]. Метод двухэтапный. На первом этапе с помощью кросс-валидации строятся два вспомогательных компонента: базовая модель отклика $\hat{m}(x) = E[Y | X = x]$ и модель propensity score $\hat{e}(x) = E[W | X = x]$. Уже на этом шаге закладывается главная особенность подхода.

В отличие от большинства конкурирующих методов, совмещающих устранение сторонних факторов и измерение эффекта воздействия в единой процедуре, R-Learner намеренно разделяет эти задачи. Подавление искажающих факторов встраивается в саму структуру функции потерь. На втором этапе остается лишь непосредственная оценка эффекта через минимизацию функционала R-loss [9]:

$$\hat{L}_n(\tau) = \frac{1}{n} \cdot \sum_{i=1}^n [(Y_i - \hat{m}(X_i)) - (W_i - \hat{e}(X_i)) \cdot \tau(X_i)]^2. \quad (14)$$

Такая конструкция позволяет выделить эффект воздействия в чистом виде, поскольку влияние среднего уровня отклика и дисбаланс между группами treatment и control оказываются устранены ещё до того, как начинается оценка $\tau(x)$.

Существенным достоинством подхода является его гибкость: на втором этапе для минимизации R-loss допускается применение любого метода машинного обучения, что позволяет адаптировать модель под конкретную задачу. Помимо этого, R-Learner устойчив к шуму в данных и не теряет качества при большом числе признаков.

Оценка качества uplift-моделей затруднена тем, что истинный эффект воздействия для конкретного клиента не наблюдаем: он не может одновременно получить коммуникацию и остаться без нее. Чтобы обойти это ограничение, субъекты упорядочиваются по величине предсказанного uplift, разбиваются на группы, и внутри каждой из них сравниваются отклики получивших воздействие с откликами тех, кто его не получил. В случаях, когда воздействие охватывает лишь часть аудитории и требуется оценить качество модели на наиболее отзывчивых клиентах, применяется метрика $uplift@k$. Она фиксирует разность средних значений целевой переменной между группами воздействия и контроля среди тех субъектов, которые модель отнесла к топ-к доле выборки по предсказанному uplift:

$$uplift@k = \bar{Y}_k^T - \bar{Y}_k^C, \quad (15)$$

где \bar{Y}_k^T и \bar{Y}_k^C – средние отклики тестовой и контрольной групп в пределах топ-k% выборки. Метрика принимает значения от $[-1; 1]$, причем значения, близкие к 1, свидетельствуют о том, что модель точно выявляет субъектов с наибольшим инкрементальным эффектом.

Когда пороговое значение k заранее неизвестно или требуется единая оценка качества модели на всей выборке, используется Weighted Average Uplift (WAU). Идея метрики состоит в том, чтобы не ограничиваться одной точкой ранжирования, а усреднить uplift по всем группам выборки, взвешивая каждую группу пропорционально числу субъектов группы воздействия в ней:

$$WAU = \frac{\sum_{b=1}^B n_b^T \cdot uplift_b}{\sum_{b=1}^B n_b^T}, \quad (16)$$

где n_b^T – число субъектов treatment группы в интервале b , а $uplift_b = \bar{Y}_b^T - \bar{Y}_b^C$ – наблюдаемый uplift внутри этого интервала. Такое взвешивание предотвращает смещение, которое возникло бы при простом усреднении по сегментам разного размера, и делает метрику устойчивой к неравномерному распределению субъектов между группами.

Для наглядной оценки способности модели концентрировать эффект воздействия в верхней части ранжированного списка строится Uplift Curve. По оси абсцисс откладывается доля субъектов, отбираемых в порядке убывания предсказанного uplift, по оси ординат – накопленный эффект воздействия:

$$UpliftCurve(t) = (\bar{Y}_t^T - \bar{Y}_t^C) \cdot (n_t^T + n_t^C), \quad (17)$$

где t – доля включенных субъектов. Чем выше кривая модели над базовой линией случайного ранжирования, тем точнее модель выделяет восприимчивых субъектов. Отношение площади между кривыми к аналогичной площади идеальной модели задает метрику Uplift AUC (Area Under Uplift Curve), которая отражает способность модели ранжировать субъектов по степени восприимчивости к воздействию.

Концептуально схожей, но методологически отличной метрикой является Qini Curve. Ее характерная особенность – штраф за дисбаланс между группами воздействия и контроля в верхней части ранжированного списка. Число откликов контрольной группы при этом нормируется пропорционально соотношению размеров групп:

$$QiniCurve(t) = n_t^T \cdot \bar{Y}_t^T - n_t^C \cdot \bar{Y}_t^C \cdot \frac{N^T}{N^C}, \quad (18)$$

где N^T и N^C – общая численность групп воздействия и контроля соответственно. Разность площадей под полученной кривой и кривой случайного ранжирования образует Qini AUC (Area Under Qini Curve). Несмотря на близость к Uplift AUC, данная метрика способна иначе упорядочивать модели – характер распределения эффектов в выборке здесь играет решающую роль. Совместное применение обеих метрик позволяет получить более полную картину.

Вместе с тем метрики ранжирования оценивают лишь одну сторону качества модели. Насколько точны сами предсказанные значения uplift, они не показывают. Этот вопрос решается с помощью Average Squared Deviation (ASD) – величины, отражающей среднее расхождение между предсказанным и фактически наблюдаемым эффектом воздействия по сегментам выборки:

$$ASD = \frac{1}{B} \sum_{b=1}^B (\hat{t}_b - \hat{u}_b)^2, \quad (19)$$

где \hat{t}_b – усредненное предсказание модели в сегменте b , $\hat{u}_b = \bar{Y}_b^T - \bar{Y}_b^C$ – фактический эффект воздействия, зафиксированный в том же сегменте. Высокое значение ASD отражает нестабильность модели и возможное переобучение, а низкое – ее способность адекватно воспроизводить реальные эффекты воздействия.

Результаты

Для сравнения моделей использовался открытый набор данных Lenta Uplift Modeling Dataset, подготовленный розничной сетью «Лента». Он содержит обезличенные поведенческие характеристики покупателей до проведения кампании, признак получения коммуникации и метку целевого действия.

Набор данных включает 687 029 наблюдений и 194 признака, описывающих демографические характеристики клиента и историю его покупательского поведения за различные периоды до проведения кампании. К ключевым признакам относятся пол и возраст клиента, тип магазина (супермаркет или гипермаркет), количество и суммы чеков в разрезе товарных групп, доля промо-товаров в корзине, коэффициенты вариации частоты визитов и среднего чека, а также история откликов на предыдущие SMS- и Viber-коммуникации. Наблюдения разделены на группу воздействия treatment (75 % выборки) и контрольную группу control (25 %).

Программная реализация исследования выполнена на языке Python в среде Jupyter Notebook.

В рамках исследования реализованы и сопоставлены пять подходов к uplift-моделированию: S-Learner, T-Learner, Class Transformation, X-Learner и R-Learner. Для каждой модели получены индивидуальные оценки uplift по клиентам, сформированы основные метрики качества, а также проведено сравнение эффективности рассмотренных подходов на обучающей выборке (Таблица 1) и на тестовой выборке (Таблица 2).

Таблица 1 – Показатели качества uplift-моделей на обучающей выборке
 Table 1 – Quality indicators of uplift models on the training sample

	Uplift@10%	Uplift@30%	Qini AUC	Uplift AUC	WAU	ASD
S-Learner	0,1219	0,0503	0,1080	0,0593	0,0075	0,0013
T-Learner	0,2179	0,0952	0,2297	0,1263	0,0075	0,0059
Class Transformation	0,1716	0,0319	0,0733	0,0446	0,0075	0,0022
X-Learner	0,0419	0,0261	0,0422	0,0232	0,0075	0,0000
R-Learner	0,0421	0,0261	0,0444	0,0245	0,0075	0,0000

Таблица 2 – Показатели качества uplift-моделей на тестовой выборке
 Table 2 – Quality indicators of uplift models on the test sample

	Uplift@10%	Uplift@30%	Qini AUC	Uplift AUC	WAU	ASD
S-Learner	0,0269	0,0159	0,0148	0,0081	0,0076	0,0000
T-Learner	0,0224	0,0148	0,0144	0,0079	0,0076	0,0002
Class Transformation	0,1791	0,0327	0,0752	0,0457	0,0076	0,0021
X-Learner	0,0113	0,0136	0,0081	0,0045	0,0076	0,0001
R-Learner	0,0134	0,0142	0,0096	0,0053	0,0076	0,0001

По результатам оценки модели показывают различную эффективность на этапах обучения и тестирования. На этапе обучения лидирующие позиции по метрикам ранжирования занимают T-Learner и S-Learner, тогда как на этапе тестирования наилучшие результаты по всем метрикам демонстрирует Class Transformation. Наименее эффективными на этапе тестирования оказались X-Learner и R-Learner.

Визуальный анализ результатов представлен на Рисунке 1 (Uplift-кривые) и Рисунке 2 (Qini-кривые).

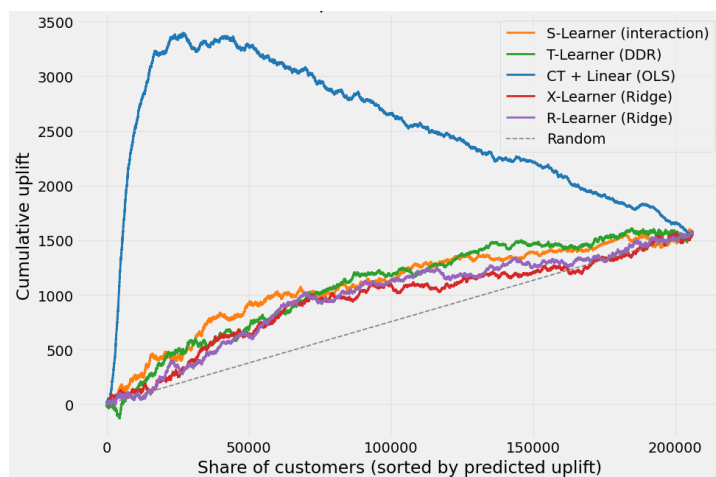


Рисунок 1 – Uplift-кривые тестовой выборки
Figure 1 – Uplift curves on the test sample

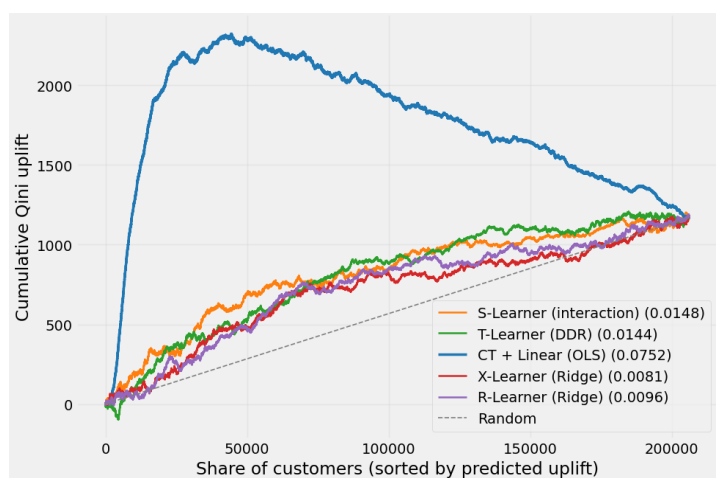


Рисунок 2 – Qini-кривые тестовой выборки
Figure 2 – Qini curves on the test sample

Кривая модели Class Transformation значительно превышает кривые остальных моделей на всем диапазоне охваченной аудитории. Все рассмотренные модели располагаются выше линии случайной модели, что графически подтверждает их способность выявлять наиболее отзывчивых клиентов при ранжировании аудитории.

Для оценки способности модели Class Transformation дифференцировать аудиторию по уровню отклика был построен график распределения фактического uplift по децилям (Рисунок 3). По оси X отображены децили клиентов, упорядоченные от наименьшего к наибольшему предсказанному uplift, по оси Y – фактический uplift, рассчитанный как разность конверсий в группах treatment и control.

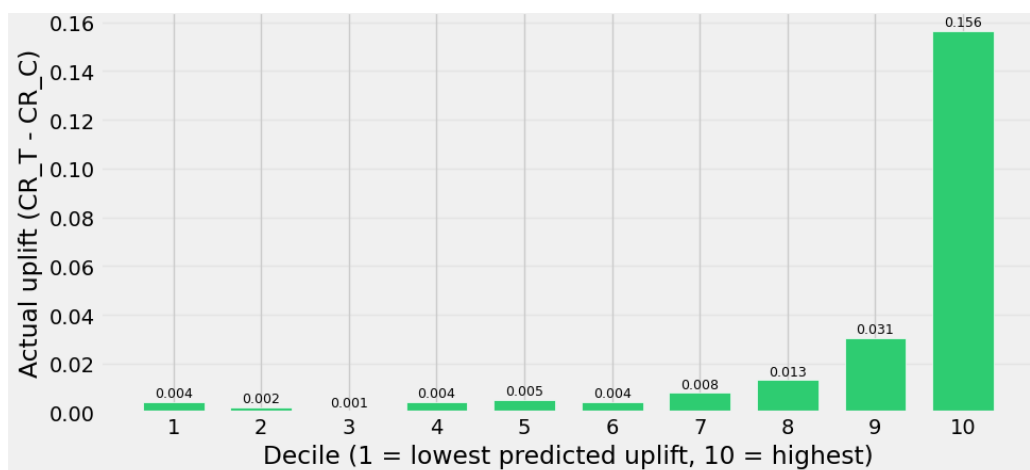


Рисунок 3 – Uplift-значения по перцентилям (Class Transformation)
Figure 3 – Uplift by percentile (Class Transformation)

Обсуждение

Полученные результаты показывают, что выбор модели uplift-моделирования существенно влияет на качество ранжирования аудитории, а высокие значения метрик на обучающей выборке не гарантируют аналогичной эффективности на тестовой. Так, в процессе обучения T-Learner достигла наибольшего значения Qini AUC, однако на тестовых данных ее результаты существенно снизились, что указывает на переобучении. Графически это проявляется в том, что кривая T-Learner на начальном участке уходит ниже нулевой отметки, а X-Learner и R-Learner на протяжении большей части диапазона практически совпадают с линией случайной модели, что соответствует их невысоким значениям Qini AUC.

Среди всех рассмотренных моделей только Class Transformation продемонстрировала наилучшее соотношение качества и стабильности. Этот результат согласуется с теоретическими преимуществами подхода: преобразование задачи в бинарную классификацию позволяет применять хорошо регуляризованные классификаторы, что снижает риск переобучения и улучшает обобщающую способность модели. Резкий подъем кривой на начальном сегменте аудитории с последующим плавным снижением визуально подтверждает концентрацию предсказательной силы в топ-децилях и приближается к форме идеальной uplift-кривой. Подобное поведение модели Class Transformation в задачах ритейл-таргетирования зафиксировано и в работах [8] и [10].

Децильный анализ позволяет оценить практическую значимость полученных результатов. Таргетирование лишь 10 % аудитории, а именно клиентов десятого дециля, обеспечивает прирост конверсии на 15,6 п.п. относительно контрольной группы. Реакция клиентов из первых восьми децилей на коммуникацию минимальна, поэтому исключение этой аудитории из кампании позволит сократить затраты, не жертвуя результатом.

Заключение

Результаты проведенного исследования показали, что не существует однозначно лучшей модели uplift-моделирования, каждая из них демонстрирует свои сильные и слабые стороны в зависимости от критерия оценки. Class Transformation оказалась наиболее подходящей для практического применения благодаря хорошей обобщенности на новых данных. Полученные выводы подтверждают, что грамотный выбор модели

способен повысить эффективность маркетинговых кампаний без роста бюджета за счет более точного определения целевой аудитории. Дальнейшие исследования могут быть направлены на тестирование ансамблевых и нейросетевых методов, анализ роли признаков отбора, а также расширение применения изученных моделей на задачи с несколькими видами взаимодействия и данные из других сфер.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Popov A., Iakovleva D. Adaptive Look-Alike Targeting in Social Networks Advertising. *Procedia Computer Science*. 2018;136:255–264. <https://doi.org/10.1016/j.procs.2018.08.264>
2. Lo V.S.Y. The true lift model: A novel data mining approach to response modeling in database marketing. *ACM SIGKDD Explorations Newsletter*. 2002;4(2):78–86. <https://doi.org/10.1145/772862.772872>
3. Hansotia B., Rukstales B. Incremental value modeling. *Journal of Interactive Marketing*. 2002;16(3):35–46. <https://doi.org/10.1002/dir.10035>
4. Rzepakowski P., Jaroszewicz S. Decision trees for uplift modeling with single and multiple treatments. *Knowledge and Information Systems*. 2012;32(2):303–327. <https://doi.org/10.1007/s10115-011-0434-0>
5. Künzel S.R., Sekhon J.S., Bickel P.J., et al. Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences of the United States of America*. 2019;116(10):4156–4165. <https://doi.org/10.1073/pnas.1804597116>
6. Devriendt F., Moldovan D., Verbeke W. A literature survey and experimental evaluation of the state-of-the-art in uplift modeling: A stepping stone toward the development of prescriptive analytics. *Big Data*. 2018;6(1):13–41. <https://doi.org/10.1089/big.2017.0104>
7. Kane K., Lo V.S.Y., Zheng J. Mining for the truly responsive customers and prospects using true-lift modeling: Comparison of new and existing methods. *Journal of Marketing Analytics*. 2014;2(4):218–238. <https://doi.org/10.1057/jma.2014.18>
8. Jaskowski M., Jaroszewicz S. Uplift modeling for clinical trial data. In: *Proceedings of the 29th International Conference on Machine Learning (ICML 2012), 26 June – 01 July 2012, Edinburgh, Scotland, UK*. 2012. URL: <https://home.ipipan.waw.pl/sj/pdf/logsemi.pdf>
9. Nie X., Wager S. Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*. 2021;108(2):299–319. <https://doi.org/10.1093/biomet/asaa076>
10. Gubela R.M., Lessmann S., Jaroszewicz S. Response transformation and profit decomposition for revenue uplift modeling. *European Journal of Operational Research*. 2020;283(2):647–661. <https://doi.org/10.1016/j.ejor.2019.11.030>

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Азарнова Татьяна Васильевна, доктор технических наук, профессор, профессор Воронежского государственного университета, Воронеж, Российская Федерация.
e-mail: ivdas92@mail.ru

Tatyana V. Azarnova, Doctor of Engineering Sciences, Professor, Professor of Voronezh State University, Voronezh, the Russian Federation.

Реброва Юлия Игоревна, студентка, Воронежский государственный университет, Воронеж, Российская Федерация.
e-mail: rebrovaiulya@mail.ru

Julia I. Rebrova, Student of Voronezh State University, Voronezh, the Russian Federation.

*Статья поступила в редакцию 09.04.2026; одобрена после рецензирования 09.06.2026;
принята к публикации 24.06.2026.*

*The article was submitted 09.04.2026; approved after reviewing 09.06.2026;
accepted for publication 24.06.2026.*