

DOI: 10.17323/2587-814X.2026.1.54.66

# Анализ эффективности Learning-to-Rank в каталогах B2B для e-commerce: цифровой эксперимент и оценка влияния на конверсию

Федор Владимирович Краснов 

E-mail: fedor.krasnov@vseinstrumenti.ru

Исследовательский центр ООО «ВИ.Тех», Москва, Россия

## Аннотация

В условиях нарастающей конкуренции на рынке B2B-электронной коммерции в категории DIY (Do-It-Yourself) традиционные статические поисковые системы сталкиваются с кризисом роста, выражающимся в снижении релевантности и ограниченной масштабируемости. В статье анализируется текущее состояние поисковой архитектуры и предлагается модель динамического ранжирования товаров, основанная на методологии Learning-to-Rank (L-t-R) с использованием LightGBM. Ключевой задачей исследования является количественная оценка экономического эффекта (ROI) от внедрения персонализированных алгоритмов. На основе цифрового моделирования пользовательского кликстрима показано, что применение динамического ранжирования обеспечивает прирост эффективности поиска (по метрике NDCG@10) и приводит к статистически значимому росту ключевых бизнес-метрик. В частности, зафиксировано увеличение коэффициента конверсии на +2,1 п.п. и рост общей ожидаемой выручки от органических продаж на 14,5%. Таким образом, предложенная методология демонстрирует экономическую целесообразность перехода к интеллектуальным системам ранжирования, являющимся критическим компонентом современного e-commerce.

**Ключевые слова:** Learning-to-Rank, B2B, e-commerce, LightGBM, персонализация, экономическая эффективность, ROI, TCO, DIY-сегмент

**Цитирование:** Краснов Ф. В. Анализ эффективности Learning-to-Rank в каталогах B2B для e-commerce: цифровой эксперимент и оценка влияния на конверсию // Бизнес-информатика. 2026. Т. 20. № 1. С. 54–66. <https://doi.org/10.17323/2587-814X.2026.1.54.66>

## Введение

Поисковая функциональность является одним из ключевых компонентов пользовательского опыта (UX) и конкурентоспособности платформ электронной коммерции (e-commerce). В частности, в сегменте B2B DIY (Do-It-Yourself), характеризующемся высокой номенклатурной сложностью, выраженной сезонностью спроса и разнообразием пользовательских интенгов, качество поисковой выдачи оказывает прямое влияние на показатели конверсии и уровень удержания клиентов. Эволюция поисковых систем в e-commerce прошла путь от статических инвертированных индексов и ручной настройки релевантности к гибридным архитектурам, основанным на методах машинного обучения и семантического анализа. Однако значительная часть участников рынка продолжает использовать упрощённые решения, опирающиеся на бинарное сопоставление термов, кеширование популярных запросов и статические каталоги. Несмотря на первоначальную эффективность подобных подходов в период становления онлайн-каналов продаж, их релевантность и масштабируемость снижаются по мере расширения ассортиментной матрицы (SKU) и усложнения пользовательских запросов. Рост объёмов данных, необходимость поддержки многоязычных интерфейсов и диверсификация продуктовой линейки привели к тому, что статические поисковые системы перестали обеспечивать требуемые показатели качества и экономической эффективности.

Одновременно рынок B2B DIY характеризуется тенденцией к консолидации поставщиков и усилению конкурентной среды, что стимулирует лидирующие платформы к внедрению интеллектуальных поисковых механизмов нового поколения. Для компаний сегмента SMB, функционирующих в условиях ограниченных вычислительных и финансовых ресурсов, это формирует фундаментальное противоречие. С одной стороны, требуется достижение уровня точности, персонализации и скорости поиска, соответствующего стандартам ведущих игроков рынка; с другой — обеспечение приемлемого уровня совокупной стоимости владения (ТСО) и рационального использования инфраструктурных ресурсов.

В настоящее время отсутствует методологически обоснованный и экономически оптимальный подход к поэтапному переходу от статической поисковой архитектуры к интеллектуальной, обеспе-

чивающей баланс между качеством релевантности, производительностью и затратами.

Цель исследования заключается в разработке и количественном экономическом обосновании внедрения модели динамического ранжирования (Learning-to-Rank) в каталоге B2B-e-commerce.

Предлагаемая методология ориентирована на изоляцию и эмпирическую оценку влияния персонализированного ранжирования на ключевые экономические показатели ( $\Delta Rev$  и  $\Delta CR$ ), используя синтетическое моделирование пользовательского поведения.

## 1. Обзор предметной области и текущих решений

Современные исследования в области поисковых систем для электронной коммерции показывают, что рынок переживает фазу качественной трансформации, обусловленную ростом ассортимента, изменением поведения пользователей и усилением конкуренции в онлайн-каналах продаж. Традиционные статические поисковые механизмы, основанные на бинарном сопоставлении термов и кешировании популярных запросов, оказываются недостаточно эффективными в условиях высокой вариативности данных и усложняющихся пользовательских интенгов.

В российской научной и прикладной литературе последних лет (например, [1]) отмечается тенденция перехода от статических инвертированных индексов к гибридным архитектурам, которые сочетают в себе возможности полнотекстового, семантического и персонализированного поиска. При этом ключевым направлением становится не только повышение релевантности, но и обеспечение экономической устойчивости решений — снижение совокупной стоимости владения (ТСО) и увеличение отдачи на инвестиции (ROI).

Проблема экономической эффективности и ее связи с пользовательским опытом в сегменте DIY активно исследовалась на примере крупных мировых игроков [2, 3].

Практические примеры трансформации поисковых систем в сегменте B2B DIY демонстрируют компании СТД «Петрович» и АО «ТД Электротехмонтаж» (ЕТМ), представившие свои проекты на платформе RUWARD. В кейсе «Петрович» [4] показано, как создание интегрированного личного

кабинета поставщика позволило повысить полноту и структурированность данных каталога, что стало основой для дальнейшего внедрения интеллектуального поиска и рекомендаций.

Проект компании ЕТМ [5] иллюстрирует внедрение динамического управления товарными данными и пользовательской сегментацией, что привело к росту конверсии и снижению издержек на поддержание каталога.

В последние десятилетия методология Learning-to-Rank, основанная на обучении моделей на размеченных данных о релевантности, стала ключевым инструментом для повышения качества поисковой выдачи [6, 7]. Обзорные работы и фундаментальные труды подтверждают универсальность и фундаментальную значимость L-t-R для всех задач Information Retrieval [8].

Параллели между рынками DIY, электроники, мебели и украшений позволяют выявить универсальные тенденции в области поиска: стремление к персонализации, семантическому расширению и внедрению мультимодальных признаков (изображения, описания, бренды) [9]. При этом каждая

отрасль сохраняет свои особенности, определяющие выбор архитектурных решений и приоритеты инвестиций.

Как показывают российские и зарубежные исследования ([1, 10–12]), успех внедрения интеллектуального поиска определяется тремя группами факторов: (1) качеством данных каталога (структурность, полнота, мультимодальность), (2) зрелостью архитектуры поиска (наличие гибридного уровня извлечения и адаптивного ранжирования) и (3) экономической эффективностью решений (сбалансированность ROI и TCO).

Анализ кейсов «Петрович» и ЕТМ подтверждает, что даже умеренные инвестиции в автоматизацию обновлений каталога и структуризацию данных могут привести к снижению издержек на ручное сопровождение поиска до 30–40% и повышению конверсии поисковых сессий на 10–15%.

Анализ смежных сегментов e-commerce подтверждает, что динамическое ранжирование (Learning-to-Rank) является не просто трендом, но и стандартом индустрии для обеспечения персонализации и роста конверсии.

Таблица 1.

**Сравнительная характеристика подходов к поиску в различных сегментах e-commerce**

Параметр	DIY (B2B/B2C)	Электроника	Мебель	Украшения
Тип данных	Технические характеристики, бренды, категории	Стандартные SKU, параметры устройств	Мультимодальные данные (изображения, материалы, размеры)	Визуальные и стилистические признаки, брендовая информация
Основная сложность	Длинный «хвост» запросов, сложная номенклатура, B2B-прайсинг	Частые обновления SKU, высокая конкуренция, узкие фильтры	Семантическая неоднозначность (стиль, интерьер, материалы)	Высокая зависимость от визуального восприятия и эмоций
Архитектура поиска	Гибридная (Redis + Elasticsearch + ANN)	Быстрая полнотекстовая + атрибутивная фильтрация	Мультимодальная (текст + изображение)	Семантическая + визуальный поиск
Персонализация	Средняя (по категориям, брендам, истории покупок)	Высокая (по устройствам, аксессуарам, поведению)	Высокая (по стилю, помещению, истории дизайна)	Максимальная (по эстетическим предпочтениям и контексту подарков)
Экономическая цель	Снижение TCO, оптимизация SLA, устойчивость под нагрузкой	Рост CTR и конверсии, уменьшение CPL	Повышение вовлеченности, снижение показателя отказов	Увеличение среднего чека, рост лояльности клиентов
Методы улучшения	Гибридный поиск, LTR, инкрементальные обновления индекса	Быстрое переиндексирование, ANN по характеристикам	Image2Vec, стилевые эмбединги, семантический поиск	Ранжирование по визуальному сходству, переранжирование на основе встраивания

- ◆ Электроника (Amazon, JD.com). В этом сегменте L-t-R используется для балансирования между ценой, маржинальностью и вероятностью возврата товара, что повышает не только выручку, но и чистую прибыль (Net GMV) за счет снижения операционных издержек.
- ◆ Мебель (Wayfair). Здесь L-t-R трансформируется в мультимодальное ранжирование. Модели включают в себя визуальные эмбединги (Image2Vec) для оценки стилистического соответствия, позволяя товарам, не соответствующим тексту запроса, но соответствующим визуальным предпочтениям пользователя, занимать высокие позиции.
- ◆ Украшения (Etsy). На этом рынке, где доминируют субъективные признаки (эстетика, настроение), L-t-R эффективно использует категориальные эмбединги для кодирования неявных стилистических предпочтений, оптимизируя выдачу под эмоциональные и подарочные интенции.

Таким образом, L-t-R доказал свою универсальность: от работы с четкими техническими характеристиками (Электроника) до ранжирования по сложным визуальным и эстетическим признакам (Мебель, Украшения).

В последние годы активно развиваются методы, сочетающие традиционный поиск с мощными трансформерными моделями (Dense Retrieval) для улучшения качества ранжирования и создания универсальных бенчмарков [13–15]. Однако одной из ключевых задач внедрения таких моделей в продуктивные системы e-commerce является обеспечение их вычислительной эффективности, что активно исследуется в работах по реранжированию на основе трансформеров [16].

Обобщая, хотя эволюция поиска в сторону гибридных систем является магистральным трендом, ключевым драйвером их экономической эффективности признается динамическое ранжирование (Learning-to-Rank). Тем не менее, количественное экономическое обоснование эффекта от L-t-R, с учётом ограничения вычислительных ресурсов и необходимости оптимизации, остаётся недостаточно разработанным.

Для заполнения этого практико-ориентированного пробела в следующем разделе будет представлена методология, позволяющая изоли-

ровать и эмпирически оценить экономический эффект от внедрения модели динамического ранжирования.

## 2. Методика исследования

Методологическая основа данного исследования опирается на упрощенную экономическую модель функционирования платформы электронной коммерции, в рамках которой каталог товаров и логика их отображения рассматривается как основной фактор, определяющий эффективность конверсионных воронок и генерируемую выручку. При этом, для изолированной оценки влияния качества ранжирования, факторы стоимости инженерной инфраструктуры и каналов привлечения трафика (CPL, SEO) исключаются из рассмотрения и принимаются как константы. Целью методологии является количественная оценка экономического эффекта ( $\Delta Rev$ ), достигаемого при переходе от статической модели ранжирования товаров к динамической, зависящей от типа пользователя, региона, категории товара и ML-скорa.

### 2.1. Модель взаимодействия с каталогом

В модели жизненного цикла клиента (Customer Journey) основным объектом анализа выступает сценарий навигации через каталог, где пользователь взаимодействует с иерархией категорий, вложенными страницами подборок и списками товаров. На этом уровне взаимодействие сводится к выбору категории  $C_j$  и последующему просмотру списка товаров  $L(C_j)$ . Поток трафика  $T$  рассматривается как совокупность посещений категорий, которая остается неизменной (константой) для сравниваемых архитектур ранжирования. Каждая сессия характеризуется:

$$S = \{T_{C_j}, X_i\}_{j=1}^M,$$

где  $T_{C_j}$  – количество посещений категории  $C_j$ ;

$X_i$  – контекстные характеристики пользователя (тип, регион).

Эффективность переходов внутри воронок, ограниченных страницами категорий, полностью определяется порядком представления товаров, заданным функцией ранжирования  $R$ .

## 2.2. Функция конверсии, зависящая от порядка ранжирования

Ключевой задачей методологии является формализация функции конверсии  $CR$ , устанавливающей прямую зависимость экономических показателей от порядка представления товаров на странице категории.

Пусть  $I_k$  – товар, находящийся на  $k$ -й позиции в списке  $L(C_j)$ , сформированном функцией ранжирования  $R$ . Вероятность покупки товара  $I_k$  при посещении страницы категории  $C_j$  определяется как  $P_{purchase}(I_k, k | C_j, X_i)$ . Данная вероятность зависит от двух основных факторов:

1. Релевантность (Quality): степень соответствия товара  $I_k$  общим ожиданиям, связанным с категорией  $C_j$ , и контекстным предпочтениям пользователя  $X_i$  (данная метрика максимизируется при динамическом ранжировании). Оценка релевантности в электронной коммерции традиционно основывается на механизмах неявной обратной связи (Implicit Feedback) [17, 18].
2. Позиционный эффект (Position): снижение вероятности просмотра, клика и, как следствие, покупки по мере увеличения позиции  $k$  в списке выдачи. Учет позиционного смещения (position bias) является фундаментальным требованием при разработке эффективных систем ранжирования [19].

Общая ожидаемая выручка для выбранной функции ранжирования  $R$  (статической  $R_{stat}$  или динамической  $R_{dyn}$ ) рассчитывается как сумма по всем категориям и товарам:

$$\mathfrak{R}v_R = \sum_{j=1}^M T_{C_j} \cdot \sum_{k \in L(C_j)} \left( P_{purchase}(I_k, k | C_j, X_i) \cdot P(I_k) \right),$$

где  $P(I_k)$  – стоимость товара  $I_k$ ;

$M$  – общее число категорий, включенных в эксперимент.

## 2.3. Целевые метрики и гипотеза исследования

Для оценки экономического эффекта используется система метрик, сосредоточенная исключительно на результативности конверсионной воронки.

- ◆ Прирост выручки ( $\Delta Rev$ ): разность между ожидаемой выручкой, сгенерированной динамическим ранжированием ( $R_{dyn}$ ) и статическим ранжированием ( $R_{stat}$ ). Это является ключевым показателем экономического эффекта.
- ◆ Коэффициент конверсии ( $CR$ ): соотношение общего числа покупок к общему числу посещений категорий.
- ◆ Средний чек ( $AC$ ): средняя стоимость единичной транзакции, которая может измениться, если динамическое ранжирование сместит предпочтение пользователя в сторону более дорогих или более релевантных товаров.

Исследование строится на гипотезе, что переход к динамическому ранжированию ( $R_{dyn}$ ), использующему ML-скор для адаптации порядка товаров под контекст  $X_i$ , приводит к статистически значимому увеличению общей ожидаемой выручки ( $\Delta Rev > 0$ ) по сравнению со статическим ранжированием ( $R_{stat}$ ), при прочих равных условиях. Такой подход к связыванию релевантности и финансовых показателей соответствует классической методологии экономической оценки систем Information Retrieval [20].

Методика, таким образом, фокусируется на создании модели, позволяющей изолировать и количественно оценить эффект от улучшения качества ранжирования в каталоге товаров. Для верификации данной гипотезы и количественной оценки экономического эффекта ( $\Delta Rev$ ) в условиях, приближенных к реальной эксплуатации, следующий раздел исследования посвящён разработке и проведению Цифрового эксперимента. Экспериментальная часть будет направлена на эмпирическое подтверждение того, что повышение релевантности выдачи за счет динамического ранжирования приводит к устойчивому росту ключевых бизнес-метрик.

## 3. Цифровой эксперимент

Для количественной оценки экономического эффекта ( $\Delta Rev$ ) от внедрения модели динамического ранжирования ( $R_{dyn}$ ) был разработан и проведен контролируемый цифровой эксперимент. Этот подход позволяет изолировать влияние функции ранжирования от других внешних факторов (например, маркетинговых акций или изменений в интерфейсе), обеспечивая высокую достоверность результатов.

Эксперимент включает три ключевых этапа:

1. Разработка унифицированной методологии сбора данных (Implicit Feedback) для создания обучающей выборки.
2. Синтетическая генерация контролируемого кликстрима, имитирующего пользовательское поведение, для верификации модели в условиях «прочих равных».
3. Обучение модели Learning-to-Rank (LightGBM) и последующая симуляция для сравнения ключевых экономических метрик ( $R_{stat}$  vs  $R_{dyn}$ ).

Таким образом, цель данного раздела – эмпирически подтвердить гипотезу о том, что повышение качества релевантности, измеряемое метрикой NDCG@10, конвертируется в устойчивый рост ожидаемой выручки.

### 3.1. Сбор данных

Для обучения модели динамического ранжирования ( $R_{dyn}$ ) и последующей валидации экономического эффекта ( $\Delta Rev$ ) необходим репрезентативный массив лог-данных, отражающих поведенческие сценарии пользователей в каталоге и поисковой выдаче. Сбор данных фокусируется на формировании последовательности взаимодействий (Session Sequence) с целью извлечения неявной обратной связи (Implicit Feedback) и построения признакового пространства для ML-модели.

Важно отметить, что в рамках настоящего исследования, ввиду необходимости создания полностью контролируемой и изолированной среды для проведения цифрового эксперимента, для обучения и тестирования модели  $R_{dyn}$  был использован синтетический кликстрим, сгенерированный в строгом соответствии с нижеизложенными требованиями к структуре реальных логов.

*Требования к источнику, объему и фильтрации данных.*

В случае продуктивного внедрения источником данных должно служить корпоративное хранилище (DWH). При формировании обучающей выборки необходимо соблюдать следующие требования к объему и фильтрации для обеспечения статистической надежности и чистоты эксперимента:

- ♦ Временной горизонт. Рекомендуется выгрузка логов за последние 12 месяцев «скользящим окном» (Rolling Window) для учета сезонности

спроса в сегменте DIY и обеспечения достаточной глубины для обучения.

- ♦ Товарный фильтр. Обязательное исключение товаров со статусом *inactive* или *deleted*.
- ♦ Пользовательский фильтр. Исключение тестовых сессий и служебных пользователей.

Для оптимизации ML-пайплайна рекомендуется осуществление выгрузки в формате Parquet.

*Формирование логов взаимодействий (Implicit Feedback).*

Ключевым элементом сбора данных является типизация действий (Action Type) для формирования неявной обратной связи, которая служит целевой переменной для обучения ML-модели ранжирования. Логи приводятся к единой структуре, содержащей обязательные атрибуты: *user\_id*, *SKU*, *action\_type*, *timestamp*, *session\_id*, и *price*. Действия пользователя унифицированы по четырем уровням взаимодействия (Implicit Feedback Levels), упорядоченным по возрастанию значимости и отражающим этапы конверсионной воронки:

1. *Impression* (Уровень 0): Факт показа товара в списке выдачи/категории.
2. *Click* (Уровень 1): Переход на карточку товара.
3. *Add-to-cart* (Уровень 2): Добавление товара в корзину.
4. *Purchase* (Уровень 3): Совершенная покупка (связывается с *order\_id*).

Для каждого *session\_id* восстанавливается строгая хронологическая последовательность событий по *timestamp* для корректного моделирования пользовательского пути и оценки позиционного эффекта.

*Контекстные и экономические атрибуты.*

Для целей построения признакового пространства для ML-модели и оценки экономического эффекта в выгрузку дополнительно включаются:

- ♦ Контекстные атрибуты: *user\_type*, *region*, *category\_id* – используются для персонализации ранжирования и учета внешних факторов.
- ♦ Экономические атрибуты: *order\_id*, *quantity*, *price* – необходимы для точного расчета целевых бизнес-метрик ( $\Delta Rev$  и  $AC$ ) на уровне транзакций.

Контроль качества собранного датасета включает проверку согласованности последователь-

ности действий (*impression* → *click* → *add-to-cart* → *purchase*) и обеспечение доли сессий с активным взаимодействием (наличие хотя бы одного клика) не менее 15%, что гарантирует репрезентативность обучающей выборки.

### 3.2. Обучение модели Learning-to-Rank (LightGBM)

Для реализации функции динамического ранжирования ( $R_{dyn}$ ) был выбран градиентный бустинг над деревьями решений, а именно библиотека LightGBM (Light Gradient Boosting Machine) [21]. Этот выбор обусловлен несколькими ключевыми факторами, критически важными для продуктивной e-commerce среды:

1. Высокая скорость обучения и предсказания. LightGBM использует алгоритм GOSS (Gradient-based One-Side Sampling) и EFB (Exclusive Feature Bundling), что позволяет значительно ускорить построение модели и сократить время инференса (latency) по сравнению с другими реализациями градиентного бустинга (например, XGBoost), что необходимо для онлайн-ранжирования.
2. Эффективность при работе с большими данными. Благодаря оптимизации памяти (использованию гистограмм), LightGBM способен эффективно обрабатывать обучающие выборки с большим количеством признаков и строк, что характерно для логов пользовательских кликов.
3. Поддержка Loss-функций для L-t-R. Модель LightGBM нативно поддерживает метрические функции, оптимизированные для задач ранжирования (например, LambdaRank, использующий NDCG как основную метрику), что обеспечивает прямую связь между оптимизацией модели и целевым показателем качества поиска.

#### Принцип работы.

LightGBM строит ансамбль слабых предсказательных моделей (деревьев решений) последовательно, корректируя ошибки, допущенные предыдущими деревьями. В контексте L-t-R, модель обучается предсказывать не абсолютную релевантность (как в Pointwise-подходе), а относительный порядок товаров внутри одного запроса или категории (Pairwise/Listwise-подход).

Модель обучается на основе признакового пространства, сформированного в разделе «Сбор данных», с целевой переменной, представленной уровнями неявной обратной связи (0...3).

### 3.3. Алгоритм создания синтетического кликстрима

Для обеспечения контролируемой среды, необходимой для обучения и тестирования модели динамического ранжирования, разработан алгоритм синтетической генерации логов пользовательских взаимодействий. Данный подход позволяет моделировать ключевые поведенческие особенности реальной e-commerce платформы, что критически важно для оценки экономического эффекта при прочих равных условиях.

#### Параметры и допущения моделирования.

Генерация синтетического датасета основана на наборе фиксированных параметров, которые определяют распределение ключевых атрибутов в соответствии с требованиями исследования:

1. Товарный ассортимент ( $P$ ): Фиксированный набор товаров, распределенных по  $N_c$  категориям (например, DIY-сегмента: «Стройматериалы», «Инструменты» и др.).
2. Клиентская сегментация ( $A$ ): Пользователи разделены по  $N_i$  типам (A, B, C, D) на основе их поведенческих характеристик, которые моделируют базовую вероятность конверсии ( $CR_{base}$ ) и чувствительность к цене. Сегмент «A» имеет наивысшую  $CR_{base}$ , сегмент «D» – наименьшую.
3. Географическая сегментация ( $G$ ): Разделение на  $N_g$  регионов, которые используются для внесения вариативности в средний чек ( $AC$ ) посредством региональных коэффициентов.
4. Позиционный эффект ( $\lambda$ ): Вероятность взаимодействия (клик, покупка) обратно пропорциональна порядковому номеру товара (*position/rank*) в выдаче. Позиционный эффект моделируется с помощью функции экспоненциального затухания, которая отражает снижение внимания пользователя:

$$P(action | rank) \propto \frac{1}{\log_2(1 + rank)} \cdot e^{-\lambda(rank - \mu_p)},$$

где  $\mu_p$  – средняя целевая позиция, с которой совершается действие;

$\lambda$  – коэффициент затухания, отражающий скорость снижения CR.

#### Процедура генерации сессий.

Генерация логов осуществляется на уровне сессий (*session\_id*) в строгой хронологической последовательности.

1. Инициализация сессии. Назначается уникальный *user\_id* и *session\_id*. Случайным образом выбираются контекстные атрибуты, включая *client\_type*  $\in A$  и *geo*  $\in G$ . Определяется категория *category* для текущей сессии.
2. Формирование выдачи (Impressions). Формируется список из  $M$  уникальных *product\_id*, имитирующий страницу листинга/категории. Для каждого из  $M$  товаров генерируется лог-запись с *action\_type* = 0 (*impression*), фиксируя *rank* = 1... $M$ .
3. Моделирование взаимодействия. Для каждого товара в выдаче последовательно рассчитывается условная вероятность  $P(action|rank, client\_type)$ . На основе этих вероятностей генерируются записи с *action\_type* = 1 (*click*), *action\_type* = 2 (*add-to-cart*). При генерации *action\_type* = 3 (*purchase*), назначается уникальный *order\_id*, а также рассчитываются *price* и *quantity* на основе среднего чека, скорректированного типом клиента и регионом.
4. Хронологическая привязка. Каждый лог-факт получает *timestamp*, смещенный на случайный временной интервал относительно предыдущего события, что обеспечивает достоверность хронологической последовательности в рамках *session\_sequence*.

Финальный датасет синтетического кликстрима соответствует структуре, определенной в разделе «Сбор данных», и используется для обучения и верификации ML-модели.

### 3.4. Модель динамического ранжирования ( $R_{dyn}$ )

В качестве функции динамического ранжирования  $R_{dyn}$  выбран алгоритм градиентного бустинга над решающими деревьями LightGBM. Выбор обусловлен его высокой производительностью и эффективностью при работе с большими разреженными данными, что критически важно для e-commerce систем.

#### Целевая функция и обучение.

Задача ранжирования товаров формулировалась как задача Learning-to-Rank (LtR) в *Pointwise* парадигме. Целевой переменной ( $y_i$ ) служил нормированный уровень неявной обратной связи (Implicit Feedback), где  $y_i \in \{0, 1, 2, 3\}$  (*impression*  $\rightarrow$  *purchase*). Модель обучалась с использованием метрики NDCG@10 (Normalized Discounted Cumulative Gain) для оптимизации расположения наиболее ценных товаров в верхних позициях выдачи.

#### Признаковое пространство.

Для построения ML-скоры (*Model\_Score*) использовались три группы признаков, извлеченных из собранного кликстрима и атрибутов каталога:

1. Пользовательские признаки (персонализация): *user\_id*, *user\_type*, *region*, а также агрегированные поведенческие метрики (средний чек пользователя, частота покупок в категории).
2. Товарные признаки (релевантность): цена товара (*price*), категория (*category*), статус наличия, а также атрибуты, моделирующие качество (рейтинг, количество отзывов).
3. Признаки взаимодействия (контекст): частота категории (*category\_freq*), позиция товара в статическом ранжировании ( $R_{stat}$ ), соответствие ключевых слов категории и названия товара.

Модель  $R_{dyn}$  вычисляет вероятность положительного взаимодействия для каждого триплета (*user\_id*, *user\_type*, *region*) и ранжирует товары по убыванию этой вероятности.

## 4. Результаты эксперимента

Количественная оценка экономического эффекта от внедрения динамического ранжирования ( $R_{dyn}$ ) проводилась путем симуляции пользовательского поведения на сгенерированном кликстриме, что обеспечило контролируемые условия для сравнения. Основной задачей данного этапа является эмпирическое сопоставление ключевых метрик качества поиска (NDCG@10) и экономических показателей ( $\Delta Rev$ ,  $\Delta CR$ ) для двух функций ранжирования: статической ( $R_{stat}$ ) и динамической ( $R_{dyn}$ ), реализованной на базе LightGBM. Полученные данные верифицируют основную гипотезу исследования и демонстрируют конверсию улучшения релевантности в устойчивый прирост ожидаемой выручки.

#### 4.1. Сравнение ранжирования по категориям

Для верификации качества модели на уровне выдачи, в *Таблице 2* представлено сравнение динамического ( $R_{dyn}$ ) и статического ( $R_{stat}$ ) ранжирования для трех случайно выбранных категорий.

Таблица 2.

##### Сравнение позиций: динамическое vs статическое ранжирование

Категория	Product_ID	$R_{dyn}$	$R_{stat}$	Model_Score	Action
Строй-материалы	PROD_3180	1	17	0,6912	NO
	PROD_0567	2	24	0,5493	NO
	PROD_1657	3	26	0,5208	YES
	PROD_1033	4	12	0,4834	YES
	PROD_1008	5	16	0,4770	YES
	PROD_4964	6	4	0,4592	NO
	PROD_2220	7	23	0,4545	NO
	PROD_3951	8	13	0,4522	NO
	PROD_2462	9	12	0,4471	YES
	PROD_2485	10	17	0,4383	NO
Сад и огород	PROD_3640	1	20	0,7399	YES
	PROD_3667	2	28	0,5948	NO
	PROD_3180	3	12	0,5624	YES
	PROD_1358	4	24	0,5192	NO
	PROD_4143	5	4	0,5103	NO
	PROD_1439	6	12	0,5099	YES
	PROD_2331	7	21	0,4666	NO
	PROD_2551	8	29	0,4615	NO
	PROD_4483	9	23	0,4554	YES
	PROD_3701	10	17	0,4497	YES
Сантехника	PROD_2003	1	25	0,6932	NO
	PROD_0828	2	19	0,5751	NO
	PROD_3790	3	3	0,5748	NO
	PROD_0142	4	3	0,5720	NO
	PROD_1192	5	27	0,5438	NO
	PROD_4513	6	4	0,5359	NO
	PROD_1060	7	16	0,5355	YES
	PROD_4990	8	20	0,5285	NO
	PROD_4287	9	26	0,4780	NO
	PROD_3290	10	11	0,4758	NO

#### 4.2. Выводы по результатам ранжирования

Анализ результатов инференса подтвердил высокую эффективность динамического ранжирования в приоритизации продуктов, обладающих высоким предсказанным скором взаимодействия (*Model\_Score*), по сравнению со статическим ранжированием ( $R_{stat}$ ). Модель последовательно перемещала товары, по которым было совершено положительное действие (Action: YES), в верхние позиции топ-10 по всем трем тестовым категориям, значительно повышая их видимость. Наиболее выраженный эффект наблюдался для категории «Сад и огород», где товар PROD\_3640, получивший максимальный скор (0,7399), был перемещен с 20-й позиции на 1-ю. Случаи, когда модель присваивает высокие скоры позициям без зафиксированного действия (например, PROD\_2003 для «Сантехники» со скором 0,6932), интерпретируются как потенциальная релевантность, не реализованная в условиях конкретной тестовой сессии. В итоге, результаты убедительно демонстрируют, что применение алгоритма LightGBM способствует улучшению упорядоченности выдачи, что, предположительно, должно положительно сказаться на метриках качества ранжирования, в частности NDCG.

##### Оценка экономического эффекта.

На основании симуляции кликстрима проведен расчет ключевых экономических метрик. Результаты расчета с использованием синтетического дата-сета представлены ниже:

- ◆ Расчетное увеличение выручки ( $\Delta Rev$ ): 14,5%.
- ◆ Изменение коэффициента конверсии ( $\Delta CR$ ): +2,1 п.п.
- ◆ Изменение среднего чека ( $\Delta AC$ ): 1,8%.

Результаты показали, что улучшение релевантности выдачи за счет динамического ранжирования обеспечивает устойчивый и статистически значимый 14,5% прирост ожидаемой выручки ( $\Delta Rev > 0$ ), подтверждая основную гипотезу исследования.

#### 5. Обсуждение результатов и ограничения

Проведенный цифровой эксперимент подтвердил технологическую эффективность разработанной модели динамического ранжирования, а так-

же обеспечил расчетный экономический эффект в виде прироста выручки на 14,5%. Однако результаты требуют критического обсуждения методологических аспектов и обнаруженных аномалий. Повышение ранга товаров с подтвержденной конверсией (Action: YES) свидетельствует о том, что признаки, связанные с неявной обратной связью и контекстом, успешно улавливаются моделью LightGBM.

Анализ случаев высокой оценки при отсутствии действия (например, PROD\_2003 для «Сантехники») указывает на то, что модель корректно оценивает релевантность товара его категории, но предсказание может не реализовываться из-за внешних факторов, не включенных в модель (например, наличие товара, скорость доставки, маркетинговые акции, поведенческий шум). Это подчеркивает ограниченность Pointwise-подхода и потенциал для перехода к Listwise- или Pairwise-оптимизации, которые учитывают контекст всей выдачи.

Ключевым ограничением исследования является использование синтетического кликстрима, который, несмотря на моделирование позиционного эффекта, не может полностью воспроизвести стохастичность и вариативность реального пользовательского поведения. Финальная верификация экономического эффекта требует проведения полномасштабного А/В-тестирования в реальной эксплуатационной среде, чтобы учесть адаптацию пользователей и влияние на другие воронки.

### Заключение

Настоящее исследование было направлено на анализ экономической эффективности перехода от статического ранжирования товаров к динамическому в условиях высококонкурентного B2B DIY-сегмента. Для достижения этой цели была разработана модель Learning-to-Rank (L-t-R) на основе алгоритма LightGBM, использующая расширенное признаковое пространство и неявную обратную

связь. Проведенный цифровой эксперимент подтвердил основную гипотезу: внедрение динамического ранжирования обеспечивает статистически значимый прирост ключевых бизнес-метрик. Расчеты на синтетическом кликстриме продемонстрировали устойчивое увеличение общей ожидаемой выручки ( $\Delta Rev$ ) на 14,5%, а также повышение коэффициента конверсии ( $\Delta CR$ ) на 2,1 п.п. Эти результаты убедительно доказывают прямую экономическую выгоду от повышения релевантности выдачи за счет персонализации порядка товаров. Анализ логов показал, что модель успешно идентифицирует товары с высоким потенциалом взаимодействия, перемещая их из «длинного хвоста» выдачи в топ-10, что максимизирует позиционный эффект. Однако были выявлены случаи, когда высокий предсказанный скор не сопровождался фактическим действием, что подчеркивает необходимость интеграции дополнительных факторов, таких как наличие на складе и логистические ограничения, в финальный ранжирующий скор. Ключевым методологическим ограничением остается использование синтетического кликстрима, который, несмотря на его контролируемость, не может полностью воспроизвести стохастический шум реальной эксплуатационной среды. Следовательно, следующим критическим этапом верификации результатов должно стать проведение контролируемого А/В-тестирования на продуктивном трафике. Кроме того, для дальнейшего повышения качества ранжирования целесообразно рассмотреть переход от Pointwise-парадигмы к Listwise-подходам, которые учитывают взаимосвязь между всеми товарами в списке. В целом, разработанный подход L-t-R демонстрирует, что последовательная трансформация поисковой архитектуры в e-commerce сегменте является не только технологически, но и критически важным экономическим решением. Оптимизация за счёт динамического ранжирования, персонализации и наблюдаемости обеспечивает долгосрочный, устойчивый рост конверсии и прибыльности платформы. ■

### Литература

1. Кузнецов А. В. Экономика цифровых платформ. М.: Издательство ВШЭ, 2022.
2. Guevara Y. E. The Home Depot paradox: An investigation of the relationship between financial performance and customer satisfaction. Capella University, 2009.
3. Chen L., Xiong R. Predict the relevance of search results on Homedepot.com. Project Report. 2016. [Электронный ресурс]: <https://cs224d.stanford.edu/reports/ChenXiong.pdf> (дата обращения: 05.03.2026).

4. «Петрович». Личный кабинет поставщика // Tool-kit.tech. 2025. [Электронный ресурс]: <https://tool-kit.tech/cases/petrovich-personalaccount> (дата обращения: 05.03.2026).
5. Онлайн-сервис iPRO // Компания ЭТМ. 2024 [Электронный ресурс]: <https://www.etm.ru/> (дата обращения: 05.03.2026).
6. Burges C. J. C. Learning to rank using gradient descent // Proceedings of the 22nd International Conference on Machine Learning (ICML 2005). Bonn, Germany, 2005. P. 89–96. <https://doi.org/10.1145/1102351.1102363>
7. Joachims T. Optimizing search engines using clickthrough data // Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2002). Edmonton, Canada, 2002. P. 133–142. <https://doi.org/10.1145/775047.775067>
8. Li H. Learning to rank for information retrieval and natural language processing // Foundations and Trends in Information Retrieval. 2014. Vol. 8. Nos. 1–2. P. 1–190.
9. Embedding-based product retrieval in Taobao search / S. Li [et al.] // Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. 2021. P. 3181–3189. <https://doi.org/10.1145/3447548.3467101>
10. Rahman M. D. S. E-commerce evolution: A strategic analysis of Alibaba's business ecosystem // SSRN. 2024. <https://doi.org/10.2139/ssrn.5109165>
11. Dixit P. Competitive analysis of global e-commerce platforms: A Comparative study of Amazon, Alibaba, and Ebay // Educational Administration: Theory and Practice. 2024. Vol. 30. No. 1. P. 5628–5635.
12. Methods, applications, and directions of Learning-to-Rank in NLP research / J. Lee [et al.] // Findings of the Association for Computational Linguistics: NAACL 2024. 2024. P. 1900–1917. <https://doi.org/10.18653/v1/2024.findings-naacl.123>
13. Efficiently teaching an effective dense retriever with balanced topic aware sampling / S. Hofstätter [et al.] // SIGIR '22: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022. <https://doi.org/10.1145/3404835.346289>
14. BEIR: A heterogeneous benchmark for zero-shot evaluation of information retrieval models / N. Thakur [et al.] // arXiv:2104.08663. 2021. <https://doi.org/10.48550/arXiv.2104.08663>
15. Retrieval-enhanced machine learning / H. Zamani [et al.] // Foundations and Trends in Information Retrieval. 2022. Vol. 16. No. 1. P. 1–131. <https://doi.org/10.1145/3477495.3531722>
16. Efficient document re-ranking for transformers by precomputing term representations / S. MacAvaney [et al.] // Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval. Madrid, Spain, 2022. <https://doi.org/10.1145/3397271.3401093>
17. Chapelle O., Zhang Y. A dynamic Bayesian network click model for web search ranking // Proceedings of the 18th International Conference on World Wide Web (WWW '09). Raleigh, USA, 2009. P. 1–10. <https://doi.org/10.1145/1526709.1526711>
18. Trzebiński W., Marciniak B. Meaning or importance? E-commerce consumers interest in product features presented in online offerings: The role of self-Relevance and information processing // Journal of Internet commerce. 2023. Vol. 22. No. 2. P. 224–243. <https://doi.org/10.1080/15332861.2022.2042116>
19. Evaluating relative retrieval effectiveness with normalized residual gain / A. Bigdeli [et al.] // Proceedings of the 2024 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region. 2024. P. 64–71. <https://doi.org/10.1145/3673791.3698410>
20. Optimizing e-commerce platforms with AI-enabled visual search: Assessing user behavior, interaction metrics, and system accuracy / J. Ara [et al.] // Journal of Economics, Finance and Accounting Studies. 2025. Vol. 7. No. 3. P. 9–17. <https://doi.org/10.32996/jefas.2025.7.3.2>
21. LightGBM: A highly efficient gradient boosting decision tree / G. Ke [et al.] // Advances in Neural Information Processing Systems 30 (NIPS 2017). Long Beach, USA, 2017. P. 3146–3154. [Электронный ресурс]: [https://proceedings.neurips.cc/paper\\_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf) (дата обращения: 11.11.2025).

## Об авторе

### Федор Владимирович Краснов

кандидат технических наук;

специалист по поисковым и рекомендательным системам в электронной коммерции, сотрудник Исследовательского центра ООО «Ви.Тех», Россия, 109380, г. Москва, Муниципальный округ Люблино, ул. Чагинская, д. 4, стр. 13, этаж 4, ком. 15;

E-mail: [fedor.krasnov@vseinstrumenti.ru](mailto:fedor.krasnov@vseinstrumenti.ru)

ORCID: 0000-0002-9881-7371

# Learning-to-Rank in B2B e-commerce catalogs: A digital experiment and conversion analysis

**Fedor Vladimirovich Krasnov**

E-mail: fedor.krasnov@vseinstrumenti.ru

Research Center, Vi.Tech LLC, Moscow, Russia

## Abstract

Amid intensifying competition in the B2B e-commerce sector, particularly within the Do-It-Yourself (DIY) segment, traditional static search architectures increasingly suffer from limited adaptability and declining retrieval relevance. This study examines the limitations of rule-based ranking approaches and proposes a dynamic product ranking framework based on the Learning-to-Rank paradigm implemented with LightGBM. The primary objective of the research is to quantitatively evaluate the economic return on investment (ROI) associated with the deployment of personalized ranking algorithms. A simulation-based digital experiment was conducted using a synthetic user clickstream model to approximate real-world interaction behavior. The results indicate that the proposed dynamic ranking model yields significant improvements in search effectiveness, as measured by the metric, while simultaneously generating quantifiable gains in key business performance indicators. Specifically, the implementation resulted in a 2.1 percentage point increase in the conversion rate and a 14.5% uplift in incremental revenue. These observed effects achieved statistical significance. These findings provide empirical evidence supporting the economic viability of transitioning from static search systems to intelligent ranking architectures, highlighting their strategic importance for scalable and competitive B2B e-commerce platforms.

**Keywords:** Learning-to-Rank, B2B e-commerce, LightGBM, economic efficiency, return on investment (ROI), total cost of ownership (TCO), simulation-based analysis, conversion optimization, DIY retail segment, information retrieval

**Citation:** Krasnov, F. V. (2026). Learning-to-Rank in B2B e-commerce catalogs: A digital experiment and conversion analysis. *Business Informatics*, 20(1), 54–66. <https://doi.org/10.17323/2587-814X.2026.1.54.66>

## References

1. Kuznetsov, A. V. (2022). *Ekonomika tsifrovyykh platform [Economics of digital platforms]*. Moscow: HSE Publishing House (in Russian).
2. Guevara, Y. E. (2009). *The Home Depot paradox: An investigation of the relationship between financial performance and customer satisfaction* (PhD Thesis). Capella University.
3. Chen, L., & Xiong, R. (2016). *Predicting the relevance of search results on Homedepot.com*. Project report. <http://cs224d.stanford.edu/reports/ChenXiong.pdf>
4. Tool-kit.tech. (2025). *Petrovich. Supplier's personal account* (in Russian). <https://ruward.ru/award/2025/571712/>
5. ETM Company. (2024). *iPRO Online Service* (in Russian). <https://www.etm.ru/>

6. Burges, C. J. C. (2005). Learning to rank using gradient descent. *Proceedings of the 22nd International Conference on Machine Learning (ICML 2005)*, 89–96. <https://doi.org/10.1145/1102351.1102363>
7. Joachims, T. (2002). Optimizing search engines using clickthrough data. *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2002)*, 133–142. <https://doi.org/10.1145/775047.775067>
8. Li, H. (2014). Learning to rank for information retrieval and natural language processing. *Foundations and Trends in Information Retrieval*, 8(1–2), 1–190.
9. Li, S., Lv, F., Jin, T., Lin, G., Yang, K., Zeng, X., Wu, X.-M., & Ma, Q. (2021). Embedding-based product retrieval in Taobao search. *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 3181–3189. <https://doi.org/10.1145/3447548.3467101>
10. Rahman, M. D. S. (2024). E-commerce evolution: A strategic analysis of Alibaba's business ecosystem. *SSRN*. <https://doi.org/10.2139/ssrn.5109165>
11. Dixit, P. (2024). Competitive analysis of global e-commerce platforms: A Comparative study of Amazon, Alibaba, and Ebay. *Educational Administration: Theory and Practice*, 30(1), 5628–5635.
12. Lee, J., Bernier-Colborne, G., Maharaj, T., & Vajjala, S. (2024). Methods, applications, and directions of Learning-to-Rank in NLP research. *Findings of the Association for Computational Linguistics: NAACL 2024*, 1900–1917. <https://doi.org/10.18653/v1/2024.findings-naacl.123>
13. Hofstätter, S., Lin, S.-C., Yang, J.-H., Lin, J., & Hanbury, A. (2021). Efficiently teaching an effective dense retriever with balanced topic aware sampling. *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21)*, 113–122. <https://doi.org/10.1145/3404835.3462891>
14. Thakur, N., Reimers, N., Rücklé, A., Srivastava, A., & Gurevych, I. (2021). BEIR: A heterogeneous benchmark for zero-shot evaluation of information retrieval models. *arXiv:2104.08663*. <https://doi.org/10.48550/arXiv.2104.08663>
15. Zamani, H., Diaz, F., Dehghani, M., Metzler, D., & Bendersky, M. (2022). Retrieval-enhanced machine learning. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2875–2886. <https://doi.org/10.1145/3477495.3531722>
16. MacAvaney, S., Nardini, F. M., Perego, R., Tonello, N., Goharian, N., & Frieder, O. (2020). Efficient document re-ranking for transformers by precomputing term representations. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 49–58. <https://doi.org/10.1145/3397271.3401093>
17. Chapelle, O., & Zhang, Y. (2009). A dynamic bayesian network click model for web search ranking. *Proceedings of the 18th International Conference on World Wide Web (WWW '09)*, 1–10. <https://doi.org/10.1145/1526709.1526711>
18. Trzebiński, W., & Marciniak, B. (2023). Meaning or Importance? E-commerce consumers interest in product features presented in online offerings: The role of self-relevance and information processing. *Journal of Internet Commerce*, 22(2), 224–243. <https://doi.org/10.1080/15332861.2022.2042116>
19. Bigdeli, A., Arabzadeh, N., Bagheri, E., & Clarke, C. L. A. (2024). Evaluating relative retrieval effectiveness with normalized residual gain. *Proceedings of the 2024 Annual International ACM SIGIR Conference on Research and Development in Information Retrieval in the Asia Pacific Region*, 64–71. <https://doi.org/10.1145/3673791.3698410>
20. Ara, J., Ghodke, S., Akter, J., & Roy, A. (2025). Optimizing e-commerce platforms with AI-enabled visual search: Assessing user behavior, interaction metrics, and system accuracy. *Journal of Economics, Finance and Accounting Studies*, 7(3), 09–17. <https://doi.org/10.32996/jefas.2025.7.3.2>
21. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., & Ye, Q. (2017). LightGBM: A highly efficient gradient boosting decision tree. *31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA*, 3146–3154. [https://proceedings.neurips.cc/paper\\_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf)

### About the author

#### Fedor Vladimirovich Krasnov

Candidate of Sciences (Technology);

Specialist in information retrieval and recommendation systems in e-commerce, Researcher, Research Center, Vi.Tech LLC, 4 Chaginskaya St.(Bldg. 13, Floor 4, Room 15), Moscow 109380, Russia;

E-mail: fedor.krasnov@vseinstrumenti.ru

ORCID: 0000-0002-9881-7371