

ЭМПИРИЧЕСКАЯ ВЕРИФИКАЦИЯ ФАКТОРОВ РАНЖИРОВАНИЯ ТОВАРОВ НА МАРКЕТПЛЕЙСЕ WILDBERRIES В КОНТЕКСТЕ МЕЖДУНАРОДНЫХ ПРАКТИК

Empirical Verification of the Factors Influencing Product Ranking on the Wildberries Marketplace in the Context of International Practices



Шарапова Виктория Олеговна,

аспирант, Саратовский государственный технический университет им. Гагарина Ю.А.; г. Саратов, Россия, ул. Политехническая, 77а

Sharapova Victoria Olegovna,

postgraduate student, Yuri Gagarin State Technical University of Saratov; 77 Politechnicheskaya St, Saratov, Russia
vika.sharapova@mail.ru

<https://orcid.org/0009-0003-1471-1858>

Цель исследования – эмпирически верифицировать факторы ранжирования товаров в поисковой выдаче российского маркетплейса Wildberries в сопоставлении с международными практиками. В статье рассматривается онтология алгоритмов ранжирования цифровых платформ, заявленные веса факторов сопоставляются с данными Ozon и Amazon, а затем проверяются на эмпирической базе из почти 3 тыс. наблюдений по 40 высокочастотным запросам одной категории за календарную неделю с помощью квартильного анализа и линейной регрессии. Научная новизна заключается в применении регрессионной модели с кластеризованными по поисковому запросу стандартными ошибками к данным поисковой выдачи российского маркетплейса Wildberries, для которого подобная эмпирическая верификация в открытой литературе ранее не проводилась. В результате установлено, что ключевыми предикторами позиции выступают количество отзывов, модель фулфилмента FBS, рейтинг, цена и процент выкупа (конверсия). Показано, что включённые факторы объясняют лишь 17,2% вариации позиции, что подтверждает доминирующую роль ненаблюдаемых параметров.

Ключевые слова: маркетплейсы; цифровая экономика; факторы ранжирования; поисковая выдача; цифровые платформы; e-commerce; конверсия; машинное обучение.

The aim of the study is to empirically verify the factors influencing the ranking of products in the search results of the Russian marketplace Wildberries, in comparison with international practices. The article examines the ontology of ranking algorithms used by digital platforms; the declared weights of the factors are compared with data from Ozon and Amazon, and then verified on an empirical basis consisting of nearly 3,000 observations for 40 high-frequency queries in one category over a calendar week, using quartile analysis and linear regression. The scientific novelty lies in applying a regression model with standard errors clustered by search query to the data from the search results of the Russian marketplace Wildberries, for which no such empirical verification has been previously conducted in the open literature. As a result, it was established that the key predictors of position are the number of reviews, the FBS fulfillment model, the rating, the price, and the purchase percentage (conversion). It has been shown that the included factors explain only 17.2% of the variation in position, which confirms the dominant role of unobserved parameters.

Keywords: marketplaces; digital economy; ranking factors; search results; digital platforms; e-commerce; conversion; machine learning.

Введение (Introduction)

Алгоритмы ранжирования товаров на маркетплейсах восходят к методам информационного поиска (*Information Retrieval*), однако существенно отличаются от классических поисковых систем по своим целевым функциям. Если традиционные поисковые системы оптимизируют релевантность файлов запросу, то маркетплейсы решают задачу, включающую еще и вероятность покупки, удовлетворённость покупателя и маржинальность для платформы. Получается, что маркетплейс преимущественно отдаёт трафик такому товару, который купят с наибольшей вероятностью, потому что именно на затратах продавцов, ассоциированных с покупкой, и зарабатывает сама платформа.

Оценка вкладов отдельных факторов в ранжирование представляет собой нестандартную методологическую задачу, обусловленную несколькими фундаментальными ограничениями. Во-первых, алгоритмы маркетплейсов в большинстве случаев являются коммерческой тайной компаний-агрегаторов и не раскрываются публично. Во-вторых, факторы ранжирования взаимосвязаны: товары с высокими продажами обычно имеют высокий рейтинг и большое количество отзывов, что создаёт проблему мультиколлинеарности при регрессионном анализе. В-третьих, поисковая выдача обновляется

в реальном времени, что делает любую оценку привязанной к конкретному периоду. И наконец, сами алгоритмы быстро обучаются и регулярно претерпевают изменения, как со стороны разработчиков, так и путем развития машинного обучения (*ML — machine learning*).

Тем не менее, понимание основ ранжирования необходимо, прежде всего, для продавцов, которым важно грамотно оценивать вклад каждого действия с карточкой товара в бизнес. Поэтому цель данной работы — анализ факторов ранжирования поисковой выдачи на российских маркетплейсах в контексте международных практик, их сравнение с иностранными маркетплейсами, а также эмпирическая верификация заявленных весов факторов на примере данных поисковой выдачи Wildberries.

Наиболее полным и ценным современным обзором алгоритмов машинного ранжирования в e-com является работа *M.A. Kabir et al* соавторами*. Авторы, часть которых являются действующими инженерами *Amazon Search*, систематизируют подходы *Learning-to-Rank (LTR, «обучение ранжированию»)*, применяемые на крупнейших мировых платформах: Amazon, Alibaba, eBay, Walmart. В работе выделены три основных класса признаков, используемых *LTR*-моделями: признаки релевантности (текстовое соответствие запроса заголовку

* A Survey on E-Commerce Learning to Rank [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/2412.03581> (дата обращения: 07.06.2026).

и описанию товара), признаки качества (рейтинг, отзывы, процент возвратов, скорость доставки) и коммерческие признаки (скорость продаж, конверсия, выручка). Ключевой вывод обзора состоит в том, что современные алгоритмы ранжирования маркетплейсов представляют собой не статические формулы с фиксированными весами, а ансамбли моделей машинного обучения, где веса признаков динамически пересчитываются в зависимости от контекста запроса, категории товара и поведенческого профиля пользователя. Таким образом, попытка определить «фиксированный вклад фактора X», по мнению авторов, является упрощением, допустимым на уровне средних тенденций, но не отражающим реальную архитектуру ML-систем ранжирования. Тем не менее, как показывает этот обзор, средние оценки вкладов факторов представляют наибольшую практическую ценность для продавцов, поскольку позволяют приоритизировать усилия по оптимизации карточек товаров.

Отдельным направлением исследований ранжирования является изучение влияния самой позиции товара в выдаче на вероятность его покупки — так называемого *position bias*. Работа G. Compiani (2024) представляет эмпирическую модель, демонстрирующую, что позиция товара в ранжированном списке является самостоятельным фактором потребительского выбора, не сводимым к характеристикам самого товара [1]. Авторы количественно оценивают «эффект позиции» и формализуют задачу оптимального ранжирования как компромисс между максимизацией выгоды потребителя и коммерческими целями платформы. Данная статья подчеркивает, что попадание в топ выдачи является не просто следствием качества товара, но и самостоятельной причиной роста продаж через механизм положительной обратной связи. Публикация J. Fong, O. R. Natan, R. Pantle (2026) также исследует поведенческий аспект восприятия ранжирования. Авторы демонстрируют, что потребители систематически делают выводы о качестве товара на основании его позиции в списке, даже когда ранжирование не отражает объективного качества**. Это создаёт ситуацию, в которой алгоритм ранжирования фактически становится инструментом формирования потребительских предпочтений, а не только их отражением.

Дополнительным вкладом в понимание роли ранжирования является работа R. Donnelly, A. Kanodia, I. Morozov (2024), в которой анализируют влияние пер-

сонализированных алгоритмов ранжирования^{3*}. Результаты показывают, что персонализация ранжирования увеличивает конверсию, но может приводить к «фильтровому пузырю», ограничивая разнообразие товаров, с которыми сталкивается покупатель. Для российских маркетплейсов это исследование тоже релевантно, по оценке MPSTATS, вес фактора персонализации в алгоритме Ozon оценивается в 11%, а в алгоритме Яндекс Маркета — до 25,5%^{4*}.

Наиболее изученным в мировой литературе является алгоритм ранжирования Amazon. Являясь самым крупным агрегатором в мире, эволюция алгоритмов Amazon хорошо прослеживается в индустриальных источниках. Классический алгоритм A9 преимущественно опирался на скорость продаж как на основной ранжирующий сигнал, тогда как обновлённый A10 использует предиктивные метрики удовлетворённости покупателя — показатели конверсии, процент возвратов и качество отзывов^{5*}. Аналитический обзор Nova Analytics (2026) документирует конкретные изменения, произошедшие в алгоритме Amazon: снижение веса внешнего трафика, усиление значимости метрик клиентской удовлетворённости и рост влияния формата фулфилмента (FBA против FBM) на позиции в выдаче^{6*}. Алгоритм A11, внедрённый на Amazon с 2024 года, делает акцент на персонализации выдачи, отдавая ей до 40% общего веса.

Стоит отметить, что значительный вес фактора продаж в алгоритмах ранжирования создаёт стимулы для манипулирования через самовыкупы, накрутку отзывов и искусственное завышение скорости продаж. Теоретико-игровая модель C. Jin, L. Yang, K. Hosanagar формализует эту проблему: авторы демонстрируют, чем выше вес скорости продаж в формуле ранжирования, тем сильнее стимул продавцов к размещению фейковых заказов, поскольку предельная выгода от повышения позиции превышает предельные затраты на накрутку^{7*}.

Китайские маркетплейсы Taobao (Alibaba Group) и JD.com демонстрируют альтернативные подходы к ранжированию, отличающиеся более выраженной ориентацией на многокритериальную оптимизацию. Исследование инженеров Taobao (2025) представляет генеративную модель переранжирования (*re-ranking*), которая учитывает не только характеристики отдельного товара, но и взаимодействие между товарами в выдаче^{8*}. Эта концепция существенно усложняет задачу эмпирической оценки вкладов отдельных факторов. Исследования Microsoft, Alibaba, JD, ByteDance и Meta

** Выводы потребителей на основе рейтингов товаров: роль убеждений в поисковом поведении — [Электронный ресурс]. — Режим доступа: https://www.researchgate.net/publication/382541592_Consumer_Inferences_from_Product_Rankings_The_Role_of_Beliefs_in_Search_Behavior (дата обращения: 07.06.2026).

^{3*} Welfare Effects of Personalized Rankings — [Электронный ресурс]. — Режим доступа: <https://pubsonline.informs.org/doi/10.1287/mksc.2023.1441> (дата обращения: 07.06.2026).

^{4*} Ранжирование карточек на Wildberries, Ozon и Яндекс Маркете [Электронный ресурс]. — Режим доступа: <https://mpstats.io/media/marketing/ranzhivovanie-na-marketpleysah> (дата обращения 01.06.2026).

^{5*} How Amazon's 2026 Algorithm Actually Ranks Products: The Factors [Электронный ресурс]. — Режим доступа: <https://amzprep.com/amazon-algorithm-ranking-factors/> (дата обращения 01.06.2026).

^{6*} Amazon Listing Optimization: A10 Algorithm Playbook: [Электронный ресурс]. — Режим доступа: <https://novadata.io/resources/blog/listing-optimization> (дата обращения 01.06.2026).

^{7*} To Brush or Not to Brush: Product Rankings, Consumer Search, and Fake Orders [Электронный ресурс]. — Режим доступа: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3468363 (дата обращения: 07.06.2026).

^{8*} Модель генеративного переранжирования для многоцелевой оптимизации на уровне списка на Taobao [Электронный ресурс]. — Режим доступа: <https://arxiv.org/html/2505.07197v1> (дата обращения: 07.06.2026).

показывают, что общая тенденция в развитии алгоритмов ранжирования — переход от линейных моделей с фиксированными весами к нейросетевым архитектурам, способным динамически адаптировать веса в зависимости от контекста запроса.

Российское научное сообщество с 2025 года активно обсуждает маркетплейсы, но преимущественно в управленческо-маркетинговом и экономическом ключе, не затрагивая вопросы эконометрики поисковой выдачи. Например, в своей работе *Балакшин И. С.* рассматривает факторы потребительского спроса, что косвенно связано с ранжированием. Автор констатирует, что покупатели «могут следовать рекомендациям алгоритмов», а также подробно раскрывает различные стимулы со стороны пользователей, в т.ч. цена, отзывы, рейтинг, репутация бренда [2].

Статья *Кружалова А.В.* — одна из немногих, где прямо затрагиваются факторы ранжирования, к которым автор относит *SEO*, качество товара, контент, отзывы, участие в акциях, динамическое ценообразование, платное продвижение. Тем не менее, данная работа составлена на уровне маркетингового обзора, а не алгоритмического исследования, и автор никак не сравнивает вклады факторов [3].

Таким образом, в российской академической литературе имеется довольно ограниченное представление об алгоритмической выдаче на отечественных маркетплейсах. Именно поэтому для составления сравнительной *таблицы 1* использованы разнородные по уровню надёжности источники: документация площадок и индустриальные аналитические материалы (*MPSTATS, AMZ Prep, Nova Analytics*) не проходят рецензирование и приводят оценки весов факторов, основываясь на опыте своих клиентов. Обращение к индустриальным источ-

никам обусловлено закрытостью алгоритмов и отсутствием количественных оценок весов факторов в открытой литературе. Эмпирическая часть настоящей работы направлена в том числе на независимую проверку этих индустриальных оценок для Wildberries.

Материалы и методы (Materials and Methods)

В литературе сложилось несколько устойчивых подходов к измерению вкладов факторов.

1. Полевой эксперимент — предполагает контролируемое изменение одного фактора (например, цены) при фиксации остальных и наблюдение за изменением позиции. Этот метод даёт наиболее чистые каузальные оценки, но требует сотрудничества с платформой или значительного бюджета на проведение, чтобы измерить влияние всех факторов.

2. Корреляционный анализ поисковой выдачи — основан на парсинге данных о позициях товаров и их характеристиках с последующим регрессионным моделированием. Данный метод является наиболее распространённым в индустриальных исследованиях и доступен внешним исследователям через аналитические сервисы. Его ограничение — невозможность разделения корреляции и причинно-следственной связи, а также проблема усечённой выборки (наблюдаемы только товары, попавшие на первые страницы выдачи).

3. Анализ естественных экспериментов — использует экзогенные шоки (изменение алгоритма, введение нового рекламного формата, сезонные акции) для идентификации каузальных эффектов. Для реализации такого метода требуются значительные временные затраты. Также этот подход сопряжен с коммерческими рисками, т.к. для проведения экспериментов нужен реальный кабинет продавца.

Таблица 1

Сравнительный анализ факторов ранжирования для Wb, Ozon и Amazon [составлено автором]
Table 1. Comparative analysis of ranking factors for Wb, Ozon, and Amazon [compiled by the author]

	Wildberries*	Ozon**	Amazon ^{3*} , ^{4*}
Конверсия	5-20%	15-25%	30-40%
Продажи / скорость продаж	10-25% объём продаж в рублях, последний месяц важнее	объединены в один фактор	10-15%
Логистика	30-40% помимо скорости доставки важен индекс локализации заказов	5-15%	8-15% (в приоритете модель фулфилмента FBA)
SEO	не указан, но спам-слова ведут к пессимизации	20-40% текстовая близость атрибутов к запросу	15-25% релевантность запроса + семантический слой
CTR / Популярность по запросу	входит в конверсию	10-15%	входит в конверсию
Акции и продвижение	7-35% (в период акции; доп. буст при совмещении с рекламой)	Бустинг на глобальных распродажах (коэффициент), платное продвижение	нет данных
Отзывы и рейтинг товара	1-10%	5-10%	до 10%
Цена	явно не указано	5-15% + бустинг за выгодный индекс цен	5-10%
Персонализация	имеет место, но без конкретного значения	10-15% (привязка к покупателю, его покупкам, интересам)	до 40% в зависимости от модели

* Выдача и ранжирование товаров на Wildberries: [Электронный ресурс]. – Режим доступа: <https://seller.wildberries.ru/instructions/ru/ru/material/item-search-results-and-ranking> (дата обращения 01.06.2026).

** Принципы работы поиска на Ozon [Электронный ресурс]. – Режим доступа: <https://docs.ozon.ru/global/promotion/product-promotions/ozon-search/?country=CN> (дата обращения 01.06.2026).

^{3*} How Amazon's 2026 Algorithm Actually Ranks Products: The Factors Nobody Talks About: – [Электронный ресурс]. – Режим доступа: <https://amzprep.com/amazon-algorithm-ranking-factors/> (дата обращения 01.06.2026).

^{4*} Amazon Listing Optimization: A10 Algorithm Playbook: – [Электронный ресурс]. – Режим доступа: <https://novadata.io/resources/blog/listing-optimization> (дата обращения 01.06.2026).

В рамках данной работы использован второй подход, для формирования эмпирической базы взят аналитический сервис *MPSTATS*. Это специализированная платформа мониторинга маркетплейсов, осуществляющая систематический сбор информации о товарах в поисковой выдаче Wildberries, Ozon и Яндекс Маркета. *MPSTATS* является одним из наиболее широко используемых инструментов аналитики российских маркетплейсов и предоставляет данные о позициях товаров, коммерческих показателях продавцов и характеристиках карточек, получаемых по *API*. Достоверность данных сервиса эксперты отрасли оценивают не менее чем на 70%.

Для исследования собраны данные по 40 поисковым запросам в рамках одной товарной категории — средства для стирки — маркетплейса Wildberries за период с 2 по 8 марта 2026 года (одна календарная неделя). Взяты самые высокочастотные запросы с частотностью не менее 10 тыс./мес; каждый кластер содержит только один запрос (самый частотный) — это сделано с целью собрать как можно более вариативную выборку без повторяющихся пресетов. Выбор одной товарной категории сделан для чистоты данных, т.к. структура весов факторов может различаться между категориями, что привело бы к неинтерпретируемым результатам.

Период наблюдения — одна неделя — выбран по двум причинам. Во-первых, недельный интервал обеспечивает достаточное сглаживание внутридневных колебаний позиций (алгоритм ранжирования Wildberries обновляет выдачу в реальном времени, не публикуя конкретный период ее «пересборки»). Во-вторых, недельный период в меру короткий, чтобы значения факторов (продажи, рейтинг, цена) и позиция товара соответствовали

одному и тому же состоянию карточки: на более длительных интервалах данные получатся слишком усредненными, и выборка не будет прецизионной.

По каждому запросу были выгружены данные по товарам с первой страницы поисковой выдачи (раздел «Товары в поиске»). Для датасета выбраны следующие метрики товаров, описанные в *таблице 2* вместе с проверяемыми гипотезами.

Все гипотезы сформулированы как предположения о направлении связи, в регрессионной модели они проверяются как наличие статистически значимого коэффициента соответствующего знака.

В качестве зависимой переменной использована средняя органическая позиция товара в поисковой выдаче за период, очищенная от эффекта платного продвижения. Меньшее значение соответствует более высокой позиции.

Набор независимых переменных (предикторов) сформирован на основании факторов ранжирования, декларируемых платформой Wildberries и подтвержденных индустриальными аналитическими сервисами как *MPSTATS*. В модель включены 12 предикторов, описанных в *таблице 2*.

Переменные «объем продаж» и «количество отзывов» логарифмированы с использованием натурального логарифма с добавлением единицы ($\ln(x+1)$) для нормализации выраженных правосторонне асимметричных распределений: например, медиана отзывов — 4 155 при максимуме 1 207 170.

Из модели намеренно исключены переменные, являющиеся производными от включенных предикторов: выручка (включает количество заказов и цену), общая позиция в выдаче (включает рекламную компоненту

Таблица 2

Метрики и проверяемые гипотезы исследования
Table 2. Metrics and tested hypotheses

	Метрика	Проверяемая гипотеза
Логистические	Количество складов, включая FBS	Более широкое складское распределение остатков повышает видимость товара в регионах и способствует повышению в выдаче
	Наличие FBS	Платформе выгоднее продвигать товары селлеров, которые пользуются логистической архитектурой маркетплейса (по FBO), т.к. это повышает доходы маркетплейса (больше заказов – больше трат продавца на логистику). Товары по FBS могут иметь заниженные позиции, т.к. минуют логистические затраты, т.е. маркетплейс с их продажи «зарабатывает» меньше. (Гипотеза опровергнута, см. результаты)
	Дней в наличии	Аутсток хотя бы на несколько часов наносит непоправимый вред карточке и полностью уничтожает ее высокие позиции
Коммерческие	Количество заказов	Чем больше объем продаж, тем больше логистические затраты селлера, тем выше должна быть позиция карточки в выдаче
	Цена с СПП, руб.	Более выгодное предложение вызывает повышенную активность пользователей, поэтому дешевые товары могут ранжироваться выше
Продвижение	Количество поисковых запросов	Из-за широкого SEO товар показывается по большему количеству ключей, что увеличивает показы и нецелевой трафик, снижая CTR. Понижение конверсий сигнализирует о пониженном спросе, что может занижать позиции
	Средний процент выкупа	Высокая конверсия на последнем этапе может косвенно способствовать повышению объема продаж (товар действительно продан, а не заказан). Повышенные возвраты не выгодны маркетплейсу: логистические мощности тратятся без экономического эффекта
	Дней на маркетплейсе	Проблема «холодного старта» должна пессимизировать карточку
	Дробный рейтинг, средний рейтинг последних 15 отзывов, количество отзывов	Высокий рейтинг и много отзывов повышают доверие покупателей и кликабельность карточки. Высокий CTR должен продвигать карточку вверх в выдаче
	Количество акций	Участие в акциях понижает цены на товары и способствует увеличению спроса (аналогично коммерческим факторам)
	Средняя позиция в органике, средняя позиция в рекламе	Внутреннее продвижение предусмотрено для повышения позиций товаров, занимающих низ выдачи. Наличие рекламы не должно оказывать влияния на органические позиции.

и, следовательно, является частично зависимой), а также дублирующий показатель рейтинга (обобщённая оценка карточки коррелирует с последними отзывами).

Анализ проведён в несколько этапов. На первом этапе рассчитана описательная статистика по всем переменным (средние, медианы, стандартные отклонения, квартили) для общей характеристики выборки.

На втором этапе проведён квартильный анализ: выборка разделена на четыре равные группы по значению органической позиции (топ-25%, 25–50%, 50–75%, аутсайдеры 75–100%), по каждой группе рассчитаны медианные значения всех факторов. Медиана выбрана вместо среднего арифметического ввиду выраженной правосторонней асимметрии распределений большинства переменных. Данный этап носит описательный характер и не предполагает статистического тестирования, однако позволяет визуально выявить структурные различия между группами.

На третьем этапе построена линейная регрессионная модель методом наименьших квадратов (МНК), в которой зависимой переменной является органическая позиция товара, а независимыми — 12 перечисленных предикторов. Стандартные ошибки коэффициентов кластеризованы на уровне поискового запроса (40 кластеров) для учёта внутригрупповой корреляции наблюдений: товары, конкурирующие за позиции в рамках одного поискового запроса, не являются независимыми наблюдениями. Кластеризация стандартных ошибок является стандартным подходом для данных с групповой структурой.

Для обеспечения сопоставимости вкладов факторов, измеренных в различных единицах (рубли, баллы, проценты, штуки), дополнительно рассчитаны стандартизированные коэффициенты (бета-коэффициенты) путём z-нормализации всех переменных модели (вычитание среднего и деление на стандартное отклонение). Стандартизированный коэффициент показывает, на сколько стандартных отклонений изменяется зависимая переменная при изменении предиктора на одно стандартное отклонение.

Проверка мультиколлинеарности предикторов проведена на основании матрицы парных корреляций Пирсона. Все вычисления выполнены с использованием языка программирования *Python 3*.

Ограничения

Исследование имеет ряд ограничений, которые необходимо учитывать при интерпретации результатов.

1. *Усечённая выборка.* Данные охватывают только товары, попавшие на первые страницы выдачи по выбранным запросам. Товары, располагающиеся на более глубоких позициях, не наблюдаемы, что создаёт проблему усечённой выборки и ограничивает экстраполяцию результатов на всю совокупность товаров.
2. *Одна товарная категория.* Результаты получены для конкретной товарной категории и могут не распространяться на другие. Структура весов факторов ранжирования, скорее всего, различается между категориями с разной конкурентной динамикой, средним чеком и поведением покупателей.

3. *Один временной срез.* Поперечная модель не позволяет установить каузальные связи и отличить причину от следствия. Панельное исследование обеспечило бы более надёжную идентификацию, но существенно усложнило бы сбор данных ввиду динамического состава выдачи.

4. *Выбранный период* (2–8 марта 2026 г.) включает предпраздничные дни. Для категории «средства для стирки» сезонный эффект минимален, т.к. категория не относится к подарочным или импульсным покупкам, традиционно ассоциированным с 8 марта. Тем не менее, не исключено общее повышение покупательской активности на маркетплейсе в этот период, которое может равномерно затрагивать всю категорию и потенциально влиять на поведенческие метрики.

5. *Отсутствующие переменные.* Конверсия, персонализация и поведенческие метрики не доступны для наблюдения в используемом источнике данных. Их пропуск может приводить к смещению коэффициентов при коррелированных предикторах. Так, коэффициенты при переменных, положительно коррелирующих с пропущенной конверсией (рейтинг, отзывы), могут быть занижены, поскольку часть их эффекта поглощается конверсией.

6. *Ограничение переменной FBS.* *MPSTATS* не позволяет выгрузить отдельные данные по количеству складов с разными моделями фулфилмента (*FBO* и *FBS*). Переменная «количество складов» отражает суммарное число складов обоих типов, а переменная *FBS* только факт наличия модели *FBS* у продавца. Тем не менее, для наиболее полной картины стоит учитывать помимо соотношения моделей поставок ещё и географию складов.

7. *Усреднение позиций.* Зависимая переменная представляет собой среднюю позицию за неделю, тогда как фактическое ранжирование обновляется ежедневно с учётом персонализации, геолокации и контекста сессии. Результаты, таким образом, отражают типичную структуру выдачи, а не её реальное состояние на конкретный момент.

8. *Корреляционная, а не каузальная интерпретация.* Все коэффициенты следует интерпретировать как условные корреляции, а не как каузальные эффекты. Например, тезис «рост рейтинга на 1 балл улучшает позицию на 6,2 пункта» означает: товары с более высоким рейтингом при прочих равных условиях, контролируемых в модели, занимают более высокие позиции — но не то, что искусственное повышение рейтинга гарантирует улучшение позиции.

Результаты и обсуждение (Results and discussion)

Первичный датасет включал 3 359 наблюдений по 40 поисковым запросам. В ходе предварительной обработки из выборки были исключены две категории наблюдений. Во-первых, 249 наблюдений (7,4%) с нулевым значением органической позиции, интерпретируемым как отсутствие товара в поисковой выдаче по данному запросу в период наблюдения (артефакт формирования отчёта *MPSTATS*). Во-вторых, 154 наблюдения (4,6%) с нулевым значением рейтинга, свидетельствующим об

отсутствии завершённых покупок с обратной связью — для таких товаров значения исследуемых факторов ранжирования (рейтинг, отзывы, процент выкупа) не определены, и их включение в модель приводило к систематическому искажению результатов.

После очистки рабочий датасет составил 2 956 наблюдений. Описательная статистика выборки представлена ниже в *таблице 3*. Квартили в данной таблице описывают распределение каждой переменной по отдельности.

Средняя органическая позиция товара в выборке составляет 109,7 (стандартное отклонение 70,1), медиана — 93, что свидетельствует о правосторонней асимметрии распределения, т.е. большинство товаров сконцентрировано в первой трети выдачи, при этом имеется «хвост» аутсайдеров с позициями вплоть до 435.

Медианная цена — 509 рублей при ценовом диапазоне от 63 до 2 336 руб. Медианное количество отзывов — 4 155, при этом среднее значение существенно выше

(30 126), что отражает наличие товаров-бестселлеров с сотнями тысяч отзывов, задающих выраженный правосторонний «хвост» распределения. Медианный рейтинг — 4,59 из 5,0.

Медианный возраст карточки составил 512 дней (почти полтора года) при разбросе от 8 (новые карточки) до 2 234 дней (более 6 лет). Медианное число складов — 2, максимальное — 155 (вероятно, есть крупный мультибрендовый продавец с распределённой логистикой).

Медианный рекламный буст (разница между органической и общей позицией) составил 13 позиций, т.е. в среднем реклама поднимает товар на 13 позиций вверх в поисковой выдаче по сравнению с его органической позицией. Минимальный рекламный буст — 144 показывает, что есть товары с неэффективной рекламой, которая пессимизирует карточку относительно ее органической позиции.

На *рисунке 1* представлена матрица парных корреляций Пирсона. Максимальная корреляция между

Таблица 3

Описательная статистика выборки
Table 3. Descriptive statistics of the sample

	заказы, шт	рейтинг последних отзывов	рейтинг карточки	количество отзывов	процент выкупа	fbс	количество складов	цена с СПП	дней в наличии	дней на сайте	количество поисковых запросов	видимость в поиске	количество акций	органическая позиция	рекламная позиция	рекламный буст
среднее	213,0	4,5	4,8	30 126,9	94,4	0,6	3,4	562,3	6,9	657,1	301,8	224,6	1,7	109,7	37,392	72,281
стандартное отклонение	481,5	0,4	0,1	103 116,8	1,1	0,5	4,0	282,4	0,8	561,6	207,5	176,1	2,0	70,1	40,9	76,0
min	0	1	4	1	87	0	0	63	0	8	1	0	0	5	0	-144
25%	0	4,33	4,8	614	93	0	1	361,75	7	193	147	86	0	58	0	13
50% (медиана)	13	4,59	4,9	4 155	95	1	2	509	7	512	275	194	1	93	28	64
75%	185	4,8	4,9	17 093	95	1	5	688	7	902	407	310	3	152	66,25	119
max	5 119	5	5	1 207 170	97	1	155	2 336	7	2 234	1 310	1 202	12	435	238	368

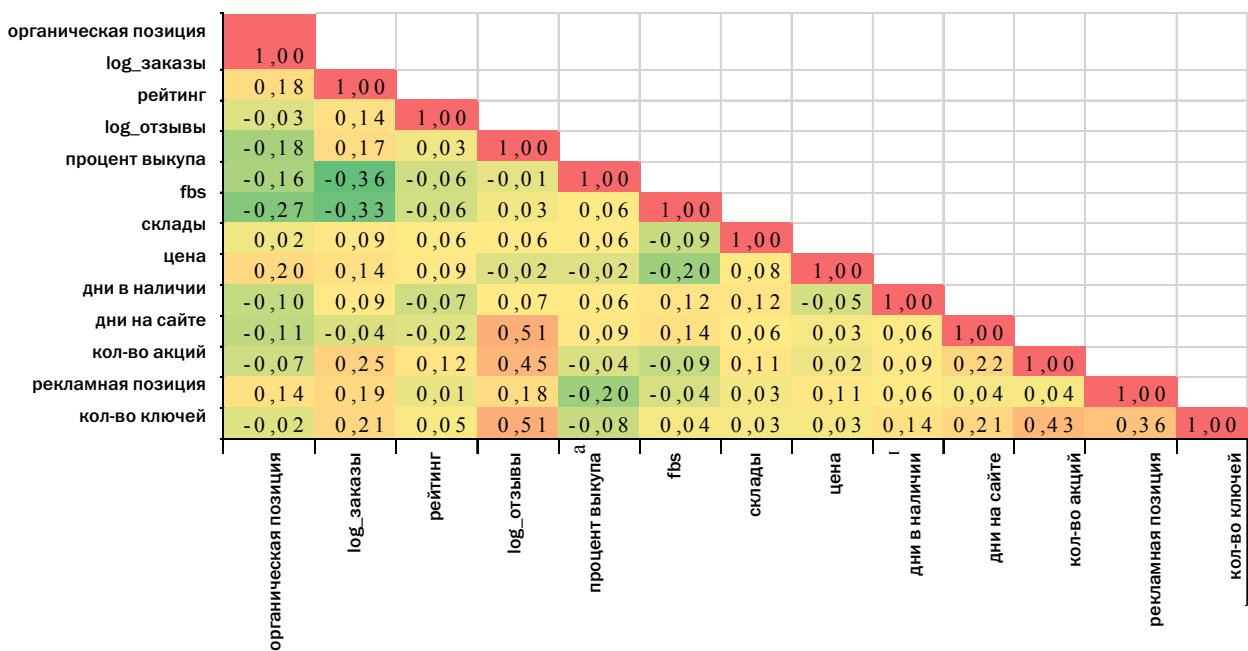


Рис. 1. Корреляционная матрица факторов ранжирования
Fig. 1. Pearson correlation matrix of ranking factors

предикторами составила 0,51 (между логарифмом числа отзывов и количеством ключевых слов), что не превышает критического порога мультиколлинеарности (0,7–0,8). Это свидетельствует о допустимости включения всех переменных в единую модель.

Для выявления общих закономерностей было произведено разделение выборки на квартили по органической позиции. Результаты квартильного анализа занесены в *таблицу 4*.

Доля FBS-товаров монотонно убывает от топ-категории (80,2%) к аутсайдерам (40,8%). Разрыв между крайними группами составляет 39,4 процентных пунктов. При этом медианное число складов в топ-квартиле – 2, в остальных группах – 3. Получается, работа по FBS не занижает ранжирование карточки, вопреки расхожему мнению экспертов отрасли.

Медианное количество отзывов в топ-квартиле (7 490) втрое превышает значение в нижнем квартиле (2 434) и также монотонно убывает от топа к низу выдачи. При этом закономерного убывания по рейтингу не наблюдается.

Медианная цена в топе – 420 руб., у аутсайдеров – 571 руб. (разрыв 26,5%). Низкая цена косвенно влияет на все этапы конверсии (дешевое предложение более привлекательно для пользователей), поэтому данный паттерн может быть устойчивым.

Медианный возраст карточки в топ-квартиле (643 дня) превышает значение в нижнем квартиле (451 день), что свидетельствует о преобладании зрелых карточек в вер-

хних позициях. Это подтверждает наличие «холодного старта» как главной проблемы новых товаров. Гипотеза о бусте новинок не нашла эмпирического подтверждения (статистически незначима).

Рекламный буст демонстрирует выраженную обратную зависимость от органической позиции: в топ-квартиле он составляет 3 позиции, тогда как в нижнем квартиле 163 позиции. Товары с высоким органическим ранжированием не просто не зависят от рекламного продвижения, а не нуждаются в нем. Аутсайдеры же компенсируют низкую позицию затратами на внутреннюю рекламу.

Медианное значение заказов в топ-квартиле составляет 0, при этом среднее – 159. Данное расхождение свидетельствует о неоднородности состава топ-квартиля, это связано с методологией оценки продаж аналитическим сервисом. С 2025 года Wildberries ограничил передачу коммерческих данных продавцов с целью умирить жалобы на конфиденциальность. Это в значительной мере усложнило сбор данных для внешних сервисов, что привело к использованию модельных оценок и предиктивных ИИ-моделей вместо фактических данных API/платформы [i].

Для оценки статистической значимости и относительного вклада каждого фактора в позицию товара при контроле за остальными наблюдаемыми характеристиками, был проведен регрессионный анализ, результаты которого представлены в *таблице 5*. Модель с 12 предикторами и кластеризованными стандартными ошибками

Таблица 4

Квартильный анализ
Table 4. Quartile analysis

	Заказы, шт	Рейтинг	Кол-во отзывов	процент выкупа	FBS	доля FBS, %	Кол-во складов	цена с СПП	дней в наличии	дней на сайте	Кол-во акций	органическая позиция	рекламная позиция	рекламный буст	заказы, шт (среднее)
Топ (1-25%)	0	4,59	7 490	95	1	80,2	2	420	7	643	2	35	28	3	158,70
25-50%	13	4,67	4 316	95	1	55,5	3	513	7	587	2	77	26	48	265,82
50-75%	36	4,67	2 974,5	94	0	46,1	3	548	7	402	1	120	0	97	257,66
Аутсайдеры (75-100%)	37,5	4,53	2 434,5	94	0	40,8	3	571	7	451	0	197	37	163	171,31

Таблица 5

Регрессионный анализ
Table 5. Regression analysis

	Нестандартизованный коэффициент	Ст. ошибка	β-коэф	p	Значимость	Доверительный интервал (ДИ)
константа	831,1434	249,2183	-	0,0009	***	-
log(заказы)	2,5023	0,9622	0,0911	0,0093	***	[0,616; 4,388]
рейтинг	-11,1599	3,7716	-0,0662	0,0031	***	[-18,552; -3,768]
log(отзывы)	-6,2559	1,1748	-0,2201	0	***	[-8,559; -3,953]
процент выкупа	-6,3849	2,7171	-0,0979	0,0188	**	[-11,710; -1,060]
FBS	-28,6468	4,5571	-0,2029	0	***	[-37,579; -19,715]
кол-во складов	0,2818	0,3974	0,016	0,4783		[-0,497; 1,061]
дни в наличии	-5,8111	1,9945	-0,0675	0,0036	***	[-9,720; -1,902]
дней на сайте	0,004	0,0032	0,0319	0,2088		[-0,002; 0,010]
цена с СПП	0,0302	0,0081	0,1216	0,0002	***	[0,014; 0,046]
кол-во поисковых запросов	0,0165	0,011	0,0489	0,1333		[-0,005; 0,038]
кол-во акций	-1,2956	0,9879	-0,0376	0,1897		[-3,232; 0,641]
рекламная позиция	0,1883	0,0515	0,1099	0,0003	***	[0,087; 0,289]

Примечание: зависимая переменная – органическая позиция, пунктов (меньшие значения соответствуют более высоким позициям выдачи). Нестандартизованные коэффициенты выражены в пунктах позиции на единицу соответствующей метрики. β-коэффициенты безразмерны и показывают изменение зависимой переменной в стандартных отклонениях при изменении предиктора на одно стандартное отклонение.

по запросу объясняет 17,2% вариации органической позиции ($R^2 = 0,172$). Восемь из двенадцати предикторов показали статистически значимую связь с зависимой переменной.

Наибольший стандартизированный коэффициент имеет количество отзывов ($\beta = -0,220$, $p < 0,001$): удвоение числа отзывов улучшает позицию товара на 6,3 пункта. Отзывы — это одновременно и прямой фактор ранжирования, и косвенный индикатор накопленного социального доказательства.

Вторым по силе влияния является наличие модели *FBS* ($\beta = -0,203$, $p < 0,001$): при прочих равных, товары с *FBS* ранжируются на 28,6 позиции выше, чем товары без *FBS*. После включения *FBS* в модель переменная «число складов» потеряла статистическую значимость ($p = 0,478$). Исходная гипотеза о пессимизации *FBS*-артикулов не подтвердилась: эмпирически *FBS* улучшает органическую позицию, а не ухудшает её. Возможное объяснение: алгоритм учитывает не только непосредственную нагрузку на склады маркетплейса, но и скорость доставки конечному покупателю, на которую поставка из регионально-го склада продавца может влиять положительно.

Финальная цена товара ($\beta = 0,122$, $p = 0,0002$) тоже показала значимый эффект, но с положительным коэффициентом: каждые 100 рублей увеличения цены ассоциированы с ухудшением позиции на 3,0 пункта.

Рекламная позиция товара также показала значимую положительную связь с органической позицией ($\beta = 0,110$, $p < 0,001$): чем ниже (больше) рекламная позиция, тем ниже (больше) органическая позиция.

Не достигли порога статистической значимости такие факторы, как количество складов ($p = 0,478$), возраст карточки ($p = 0,209$), количество ключевых слов ($p = 0,133$) и участие в акциях ($p = 0,190$).

Наглядно сравнение β -коэффициентов продемонстрировано на рисунке 2.

Полученные результаты частично подтверждают, а частично корректируют оценки весов факторов ранжирования, представленные в индустриальных источниках.

Так, шесть из двенадцати исследованных факторов показали статистически значимую связь с органической позицией в ожидаемом направлении. Наибольшие по абсолютной величине стандартизированные коэффициенты получены для количества отзывов ($\beta = -0,220$, 95% ДИ [-8,559; -3,953] и наличия модели *FBS* ($\beta = -0,203$, 95% ДИ [-37,579; -19,715]). Близость значений этих коэффициентов не позволяет однозначно утверждать о статистически значимом различии их вкладов, однако оба фактора существенно превосходят по силе влияния остальные предикторы модели. Примечательно, что большее количество отзывов одновременно положительно влияет и на конверсию, поскольку является социальным доказательством качества и популярности товара. Анало-

гично работает и низкая цена — ее влияние не указано напрямую, но в ходе исследования выяснилось, что товары с низкой ценой занимают верхние позиции выдачи, скорее всего, по той же причине (товары в топе стоят в среднем на 26,5% дешевле аутсайдеров).

Скорость логистики, а именно наличие модели *FBS* подтвердилось как второй по силе фактор, вклад которого поглощает вклад количества складов. При прочих равных товары с *FBS* ранжируются на 28,6 позиции выше, несмотря на стойкое предубеждение продавцов в пользу *FBO*. Также подтверждены рейтинг товара, процент выкупа (один из этапов конверсии), стабильность наличия на складе. Контрольная переменная «возраст карточки» оказалась незначимой ($p = 0,209$), что опровергает гипотезу о «бусте новинок». Квартильный анализ показывает, что зрелые карточки занимают более высокие позиции (медиана 643 дня в топе против 451 дня у аутсайдеров). Это подтверждает проблему «холодного старта»: новые товары без накопленных отзывов и конверсий не получают высоких органических позиций.

Два фактора показали значимую связь с позицией, однако их содержание остаётся неоднозначным. Количество ключей в *SEO* не достигло порога статистической значимости как предиктор позиции, т.к. текстовая релевантность определяет попадание товара в выдачу по конкретному запросу, а не место внутри неё. Рекламная позиция обнаружила значимую связь с органической позицией, однако невозможно разграничить прямое влияние рекламы на органику. Вероятнее всего обратное — у товаров с сильной органикой рекламный буст «отталкивается» от нее, а не с конца выдачи, что позволяет достичь более высоких мест. К тому же, квартильный анализ показывает, что медианный рекламный буст у топ-товаров незначителен по сравнению с аутсайдерами, что косвенно свидетельствует о вторичности рекламы по отношению к органике. Участие в акциях не показало самостоятельного эффекта, возможно, их влияние реализуется через ценовую переменную, уже включенную в модель.

Единственный фактор, показавший парадоксальный паттерн — объём продаж ($\beta = +0,091$, $p = 0,009$). Индустриальные источники и сама платформа единогласно называют объём продаж главным фактором ранжирования

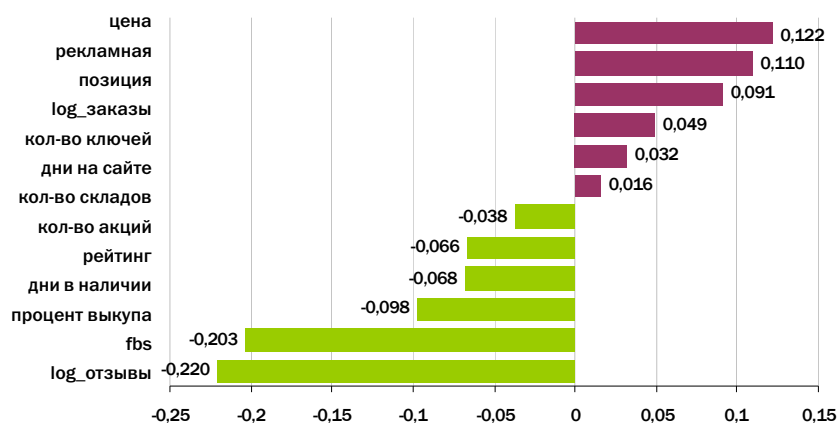


Рис.2. Стандартизированные β -коэффициенты регрессионной модели
Fig.2. Standardized regression coefficients (β)

и ожидаемым направлением «больше продаж — выше позиция». Наблюдаемое расхождение объясняется уже упомянутыми ограничениями измерения. Переменная «заказы» в *MPSTATS* (и аналогах) отражает совокупные продажи товара по всем зонам показов, тогда как алгоритм ранжирования учитывает продажи именно через поисковую выдачу. Товар может быть лидером продаж в рекомендательных полках, при этом занимая слабые позиции в поисковой выдаче по конкретному запросу. Переменная сохранена в модели непосредственно для валидации опубликованных маркетплейсом данных о факторах ранжирования, а полученный результат указывает на ограничение доступных внешних данных, обусловленное закрытостью *API* платформы^{9*}.

В совокупности результаты свидетельствуют о том, что большинство факторов ранжирования Wildberries, опубликованных самой платформой, находят эмпирическое подтверждение на наблюдаемых данных. Включённые в модель предикторы объясняют 17,2% вариации органической позиции; оставшиеся ~83% определяются принципиально ненаблюдаемыми факторами, например, персонализацией, другими этапами конверсии и поведенческими метриками сессии, доступными только самой платформе.

Заключение (Conclusion)

Проведённое исследование показало, что большинство опубликованных Wildberries факторов ранжирования находят эмпирическое подтверждение на наблюдаемых данных. Ключевыми предикторами органической позиции оказались количество отзывов, модель фулфилмента, рейтинг, цена, процент выкупа и бесперебойное наличие. Тем не менее, 83% вариации позиции определяется факторами, недоступными внешнему исследо-

вателю, что отражает механизм современных алгоритмов ранжирования — они представляют собой динамически обучаемые *ML*-системы, а не статические формулы с фиксированными весами.

Именно это обстоятельство задаёт наиболее перспективные направления дальнейшей работы. Переход к панельному дизайну позволит наблюдать, как изменение отдельного фактора, например, смена модели фулфилмента, отражается на позиции. Сравнение алгоритмов Wildberries, Ozon и прочих маркетплейсов на сопоставимых данных остаётся практически неисследованной областью, несмотря на очевидные различия в заявляемых факторах.

Наиболее перспективным направлением также представляется изучение влияния генеративного поиска на архитектуру ранжирования. Amazon уже внедрил ИИ-ассистента *Rufus*, переводящего поиск из режима «список по ключевому слову» в режим диалога с контекстом^{10*}. Wildberries и Ozon движутся в том же направлении. Это означает, что текстовые факторы ранжирования могут принципиально изменить свою роль: на смену релевантности запросу придёт релевантность покупательскому намерению, извлекаемому из диалога^{11*}.

Отдельного внимания заслуживает проблема алгоритмической справедливости. Остаётся неизученным, получают ли крупные продавцы структурные преимущества в ранжировании в силу объёма продаж, независимо от качества карточки и потребительского опыта. Если алгоритм систематически вознаграждает размер, а не качество, это создаёт барьеры входа для малого бизнеса и концентрирует рынок вокруг нескольких игроков. На англоязычных платформах эта проблема уже стала предметом академических исследований и антимонопольного регулирования [4].

ИСТОЧНИКИ (References):

1. Online Search and Optimal Product Rankings: An Empirical Framework / G. Compiani, G. Lewis, S. Peng, P. Wang // *Marketing Science*. - 2023. - V. 43. - №3. - DOI 10.1287/mksc.2022.0071.
2. Балакшин, И. С. Анализ потребительского поведения на универсальных и нишевых маркетплейсах: особенности и ключевые факторы спроса на российском рынке / И. С. Балакшин // *Вестник евразийской науки*. - 2025. - Т. 17, № 2. - DOI 10.15862/19ECVN225.
3. Кружалов, А. В. Маркетинговые стратегии конкуренции торговых компаний на цифровых площадках / А. В. Кружалов // *Прикладные экономические исследования*. - 2025. - № 6. - С. 150-155. - DOI 10.47576/2949-1908.2025.6.6.018.
4. Krishnan V. Privacy-Preserving Fair Benchmarking in Platform-Mediated Markets / V. Krishnan // *Journal of Information Systems Engineering & Management*. - 2026. - V. 11. - №1s. - P. 557-568. - DOI 10.52783/jisem.v11i1s.14137.

Поступила в редакцию / Received 20.06.2026

После доработки / Revision 27.07.2026

Принята к публикации / Accepted 11.08.2026

^{9*} Цифры любят тишину: — [Электронный ресурс]. — Режим доступа: <https://www.kommersant.ru/doc/8079374> (дата обращения 01.06.2026).

^{10*} Amazon's next-gen AI assistant for shopping is now even smarter, more capable, and more helpful [Электронный ресурс]. — Режим доступа: <https://www.aboutamazon.com/news/retail/amazon-rufus-ai-assistant-personalized-shopping-features> (дата обращения 01.06.2026).

^{11*} The Rise of AI Search: Implications for Information Markets and Human Judgement at Scale [Электронный ресурс]. — Режим доступа: <https://arxiv.org/abs/2602.13415v2> (дата обращения: 07.06.2026).