

**ПРОЕКТИРОВАНИЕ ИНТЕЛЛЕКТУАЛЬНОГО
РЕКОМЕНДАТЕЛЬНОГО ЯДРА В СИСТЕМЕ ПОДБОРА
АВТОЗАПЧАСТЕЙ**

Рассматривается проектирование интеллектуального рекомендательного ядра, предназначенного для поддержки принятия решений при подборе запасных частей для грузового автотранспорта. Актуальность разработки обусловлена усложнением необходимостью адаптации к изменяющемуся рынку запчастей и потребностью в персонализированном обслуживании клиентов. Основное внимание уделено построению алгоритмической модели ранжирования товарных позиций с учётом технических, поведенческих и логистических признаков. В качестве метода используется градиентный бустинг LightGBM. Для обучения модели формируется выборка на основе истории заказов. Проведена классификация признаков, выполнена архитектурная декомпозиция ядра и определены требования к адаптивности и масштабируемости модуля. Качество модели оценивается метриками ранжирования, включая NDCG и Precision@k. Полученные результаты демонстрируют применимость предложенного подхода для задач подбора автозапчастей в условиях нестабильных поставок и расширяющегося ассортимента.

The paper addresses the design of an intelligent recommendation core intended to support decision-making in the selection of spare parts for heavy-duty vehicles. The relevance of this development is driven by the need to adapt to a changing spare parts market and the growing demand for personalized customer service. The main focus is on building an algorithmic model for ranking product items, taking into account technical, behavioral, and logistical features. Gradient

boosting using LightGBM is applied as the core method. The model is trained on historical order data, with the training set formed using query-based grouping. Feature classification is performed, the architecture of the core is decomposed, and requirements for the adaptability and scalability of the module are defined. The quality of the model is evaluated using ranking metrics, including NDCG and Precision@k. The results obtained demonstrate the applicability of the proposed approach to the task of spare parts selection under conditions of unstable supply and an expanding product assortment.

Ключевые слова: интеллектуальное рекомендательное ядро, подбор автозапчастей, ранжирование, LightGBM, LambdaRank, поведенческие признаки, персонализация, метрики качества

Keywords: intelligent recommendation core, spare parts selection, ranking, LightGBM, LambdaRank, behavioral features, personalization, quality metrics

Современные российские компании, работающие в сфере обслуживания грузового транспорта, сталкиваются с ростом номенклатуры запасных частей, усложнением технических требований и нестабильностью логистических цепочек. Это особенно заметно в условиях постсанкционной экономики, где часть оригинальных компонентов стала недоступна, а подбор аналогов требует экспертного участия и временных затрат. В этих условиях возрастает потребность в интеллектуальных системах, способных автоматизировать подбор запчастей с учётом технической совместимости, поведения клиентов и текущей логистической ситуации [2].

Одним из решений такой задачи является проектирование интеллектуального рекомендательного ядра, встроенного в информационную систему подбора и продажи автозапчастей. Ядро должно не только учитывать параметры транспортного средства (VIN-код, тип двигателя, год выпуска), но и адаптироваться к текущей товарной доступности, срокам поставки, индивидуальным предпочтениям клиента и истории его заказов. Это особенно

актуально в сегменте грузового автотранспорта, где сбои в подборе ведут к прямым убыткам из-за простоя техники.

Информационные системы подбора запчастей традиционно основываются на фильтрации по каталогам или VIN-кодам. Однако такие подходы имеют ограничения: они не учитывают вероятностные предпочтения клиента, не адаптируются к изменениям в логистике и не предлагают структурированный набор заменителей при отсутствии оригинальной позиции [1]. В результате заказчику требуется дополнительное время на поиск подходящей альтернативы, а менеджеры перегружены ручной проверкой совместимости и сроков поставки.

Интеллектуальное рекомендательное ядро проектируется как модуль, интегрируемый в информационную систему и реализующий задачу ранжирования доступных запчастей. В отличие от простого фильтра, такой модуль формирует персонализированный список рекомендаций, в котором приоритеты определяются на основе обученной модели. Алгоритм принимает во внимание технические, поведенческие и логистические признаки, а результатом становится ранжированный список позиций, среди которых могут присутствовать как оригинальные детали, так и функционально совместимые аналоги.

Проектирование такого ядра требует формализации типов признаков, выбора устойчивой к логистическим колебаниям модели, построения обучающей выборки на основе истории заказов и определения архитектуры модуля как части системы. Задача ранжирования — наиболее подходящая, поскольку она позволяет не просто классифицировать детали как подходящие или неподходящие, а упорядочить их в соответствии с релевантностью, а это особенно важно в условиях множества допустимых вариантов [3].

Гипотеза, положенная в основу статьи, заключается в том, что интеллектуальное рекомендательное ядро, использующее задачу ранжирования и обученное на данных о заказах, поведении клиентов и логистике, способно повысить эффективность подбора запчастей. В условиях

нестабильных поставок и необходимости автоматического подбора аналогов такое ядро позволяет не только ускорить процесс выбора, но и повысить точность рекомендаций.

Цель статьи — разработать подход к проектированию интеллектуального рекомендательного ядра для информационной системы подбора автозапчастей. Для достижения этой цели решаются следующие задачи: описание архитектуры ядра и его роли в системе, классификация признаков, обоснование выбора модели ранжирования, логика формирования обучающей выборки, а также формулировка требований к модулю с точки зрения масштабируемости и адаптивности к рыночным условиям.

Актуальность выбранного подхода обусловлена не только экономической необходимостью ускорения подбора и сокращения нагрузки на сотрудников, но и внешними обстоятельствами. В условиях санкций значительная часть оригинальных запчастей становится недоступной для российских компаний, что делает задачу автоматического подбора аналогов особенно важной. Интеллектуальные модули позволяют компенсировать этот фактор за счёт анализа технической совместимости, поведения клиентов и адаптации к текущим условиям поставок. Это повышает устойчивость бизнес-процессов в условиях неопределённости.

Интеллектуальное рекомендательное ядро проектируется как часть информационной системы, автоматизирующей подбор автозапчастей для грузового транспорта. Основной задачей ядра является формирование ранжированного списка товарных позиций, соответствующих условиям конкретного запроса. В отличие от фильтрационных подходов, где результатом является набор совпадающих по параметрам товаров, в этом случае требуется упорядочивание позиций с учётом множества факторов: технической совместимости, поведения клиента, логистических ограничений и складской доступности. Постановка задачи производится в форме ранжирования (Learning to Rank), что означает, что модель обучается не на отдельные объекты, а на группы объектов (query groups), которые

соответствуют конкретным пользовательским запросам. Цель модели — определить такой порядок элементов внутри группы, при котором релевантные позиции располагаются выше нерелевантных. Этот формат уместен при наличии множества допустимых аналогов, а также при наличии неоднозначных пользовательских предпочтений.

В качестве алгоритма выбрана модель LightGBM, показавшая высокую эффективность при работе с табличными данными [4], с активацией режима LambdaRank. LightGBM показал высокую эффективность при обучении на табличных бизнес-данных, обладает встроенной поддержкой группового ранжирования и предоставляет удобные средства контроля переобучения. Метод LambdaRank позволяет учитывать относительные приоритеты товаров внутри одного заказа, что делает модель применимой к персонализированным рекомендациям, где важна не только точность попадания, но и порядок выдачи.

Обучающая выборка формируется на основе истории заказов, при этом в выборке каждая строка представляет отдельный товар, предложенный клиенту или доступный для подбора в момент оформления заказа. Отдельная заказная сессия рассматривается как query-группа [5]. Внутри группы задаются метки релевантности: так позициям, которые были куплены, присваивается максимальный приоритет, а остальным — нулевой или пониженный. Рисунок 1 иллюстрирует поэтапное формирование обучающей выборки. На начальных этапах отбираются заказы с известным исходом, формируются группы, присваиваются метки релевантности. Далее выполняется объединение с признаками из различных источников, в результате чего формируется итоговая структура, пригодная для обучения модели ранжирования.

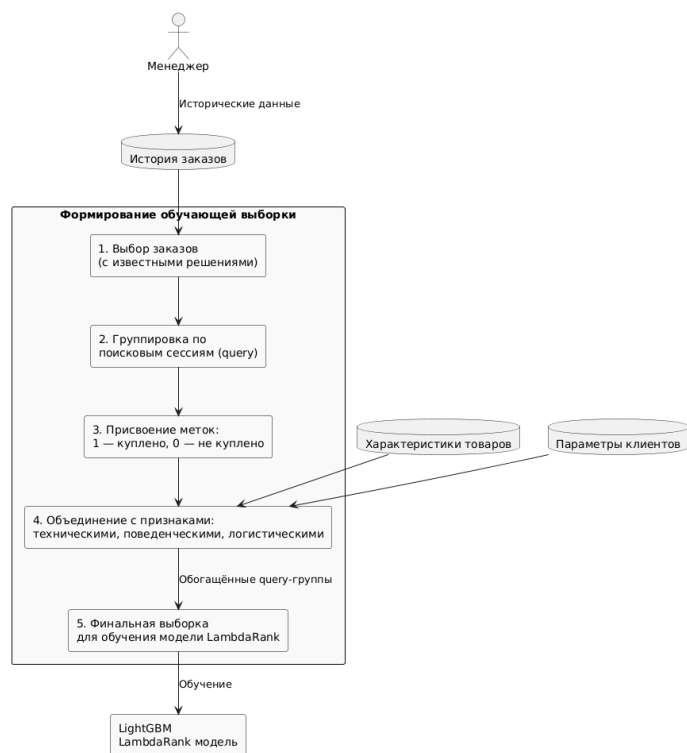


Рисунок 1 - поэтапное формирование обучающей выборки

Данные для обучения извлекались из корпоративной базы заказов за двухлетний период. Объём составляет порядка 120 000 строк после очистки и фильтрации. Структура таблицы включает: ID заказа, ID клиента, ID товара, дату обращения, VIN-код, год выпуска транспортного средства, марку, категорию запчастей, бренд товара, наличие на складе, срок поставки, поведение клиента (частота обращений и любимые бренды), цену товара, оборачиваемость, дату предыдущей покупки. Также отдельным столбцом задаётся бинарная метка факта покупки.

Все признаки делятся на три логические группы. К техническим относятся свойства автомобиля и запчастей: тип двигателя, марка, модель, VIN-код, категория узла. К поведенческим: индивидуальные предпочтения клиента, частотность заказов, средний чек, история возвратов, повторяемость брендов. К логистическим: наличие на складе, срок поставки, скорость обработки, сезонные факторы. Категориальные признаки кодируются как тип category, числовые подвергаются лог-преобразованию, нормализации и проверке на выбросы. Перед обучением исключаются пустые и противоречивые значения.

Обучение проводится в пропорции 80:20 между обучающей и валидационной выборками. В процессе обучения контролируется значение метрики $NDCG@10$ — основной показатель качества ранжирования, применяемый в рекомендательных системах. Эта метрика позволяет учитывать позицию релевантного товара в списке, что критично в условиях, при которых система должна не просто угадать нужную позицию, а поместить её как можно выше. Вспомогательно отслеживаются MAP и $Precision@5$. Обучение останавливается по критерию отсутствия улучшений в течение 30 итераций.

Для поддержания актуальности модели применяется регулярное переобучение. Обновление ядра выполняется каждые три недели на новых накопленных данных. Обучение осуществляется офлайн, без нагрузки на основной сервис. Готовая модель загружается в систему через механизм горячей замены, позволяющий переключить версию модели без перерыва в работе. Такой подход обеспечивает адаптацию к изменениям ассортимента, сезонному спросу, новым условиям поставки и поведению клиентов.

Предложенная схема проектирования и обучения позволяет воспроизвести модель в других контекстах при наличии схожих данных: заказов, информации о клиентах и складах, истории покупок. Каждый шаг — от формирования признаков до настройки модели — задаётся явно и может быть повторён при масштабировании системы или переносе в другой регион.

Интеллектуальное рекомендательное ядро было спроектировано как отдельный модуль, интегрируемый в архитектуру информационной системы подбора запчастей. Система предполагает взаимодействие нескольких компонентов: модуля агрегации данных, блока предобработки признаков, модели ранжирования и сервиса рекомендаций. Основной поток данных организован таким образом, чтобы входящие параметры запроса (например, VIN-код, категория узла, поведенческий профиль клиента) обрабатывались последовательно и передавались на вход модели для расчёта релевантности товарных позиций.

Архитектура ядра представлена в виде четырёх основных блоков. Первый — модуль извлечения данных. Он получает информацию из внутренних хранилищ компании: историю заказов, характеристики автомобилей, складские данные и справочники. Второй — модуль обработки признаков. Здесь данные нормализуются, категоризируются, дополняются производными признаками, такими как региональная оборачиваемость или брендовая повторяемость. Третий блок — модель ранжирования, реализованная на базе LightGBM. Четвёртый — REST-сервис рекомендаций, предоставляющий результаты для пользовательского интерфейса и других систем. Общая архитектура интеллектуального рекомендательного ядра представлена на Рисунке 2. Взаимодействие между модулями осуществляется через формализованные потоки данных, включая обращение к хранилищам и формирование обучающей выборки.

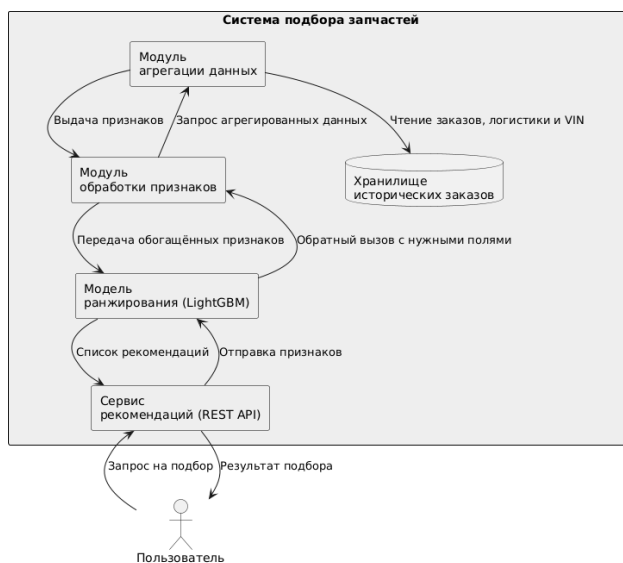


Рисунок 2 - архитектура интеллектуального рекомендательного ядра
 Признаки, используемые в модели, систематизированы в Таблице 1.

Группа признаков	Примеры признаков	Типы данных
Технические	VIN-код, тип двигателя, марка автомобиля, модель автомобиля, год выпуска, категория узла	Категориальные

Поведенческие	Любимые бренды, история заказов, доля повторных покупок, средний чек, среднее количество позиций в заказе, среднее время отклика, длительность сессии	Категориальные, числовые
Логистические	Наличие на складе, среднее время поставки, частота возвратов, региональная оборачиваемость, популярность позиции	Числовые

Таблица 1 - Признаки, используемые в модели

Всего выделено три основные группы: технические, поведенческие и логистические. К техническим отнесены: VIN-код, тип двигателя, марка автомобиля, модель автомобиля, год выпуска, категория узла. Поведенческие включают: любимые бренды, историю заказов, долю повторных покупок, средний чек, среднее количество позиций в заказе, среднее время отклика, длительность сессии. Логистические параметры охватывают: наличие на складе, среднее время поставки, частота возвратов, региональная оборачиваемость, популярность позиции. Все признаки проходят этап согласования типов, приведения к допустимому диапазону и устранения некорректных значений.

Формирование обучающей выборки выполнялось по определенной логике. Каждому заказу сопоставлялся набор потенциальных товарных позиций, доступных в момент оформления. Эти позиции формировали query-группу. Метка релевантности присваивалась на основе факта покупки: купленные товары получали максимальный приоритет. Если клиент отказывался от заказа или менял выбор в ходе общения с менеджером, это также учитывалось — метка релевантности понижалась или обнулялась. Структура группы сохранялась в процессе обучения и применялась на этапе расчёта метрик. Обучение модели проводилось на подготовленной выборке с использованием стандартной процедуры валидации. После завершения

обучения были получены следующие метрики: NDCG@10 — 0.87, MAP — 0.61, Precision@5 — 0.69. Наиболее показательным значением является NDCG, так как оно отражает не только попадание нужной позиции в выдачу, но и её ранжированную позицию. Поведение модели валидационно стабильно, расхождение с результатами на обучающем наборе не превышает допустимого порога. Это свидетельствует об адекватной обобщающей способности ядра.

Полученные результаты подтвердили работоспособность проектной архитектуры ядра. Построенная модель показала высокую согласованность между логикой рекомендаций и действиями пользователей. Использование логистических и поведенческих признаков положительно повлияло на точность модели по сравнению с вариантом, основанным исключительно на технических параметрах. Архитектура рекомендательного ядра может быть масштабирована и адаптирована под другие товарные сегменты или регионы при наличии аналогичных данных.

В статье была выполнена постановка задачи проектирования интеллектуального рекомендательного ядра, предназначенного для системы подбора автозапчастей в условиях высокой вариативности ассортимента и нестабильности поставок. В ходе работы проанализированы требования к системе рекомендаций, включая необходимость учитывать не только технические параметры транспорта, но и поведенческие особенности клиентов, а также логистические ограничения, характерные для рынка запчастей грузового транспорта. Сделан акцент на устойчивости архитектуры к внешним изменениям, включая перебои в поставках оригинальных товаров и необходимость подбора аналогов.

Выбранный подход ранжирования на базе LightGBM и метода LambdaRank доказал свою применимость для данной задачи. Метрики качества показали, что модель способна формировать релевантные рекомендации с учётом бизнес-ограничений. При этом удалось избежать типичной проблемы чрезмерной зависимости от исторических предпочтений: использование смешанных признаков позволило формировать более

сбалансированные и устойчивые предложения, даже в условиях появления новых товаров или временного отсутствия старых.

Критическая оценка результатов показала, что проектное решение устойчиво к изменениям пользовательских паттернов. Валидационные метрики оставались стабильными при тестировании модели на новых данных. Также подтверждена применимость архитектуры к сценариям расширения: ядро может быть использовано не только в рамках одной товарной группы, но и перенесено в другие сегменты с минимальной адаптацией признакового пространства.

В условиях санкционных ограничений и необходимости гибкого подбора аналогов, интеллектуальное рекомендательное ядро приобретает прикладную значимость. Использование подхода ранжирования позволяет быстро перестраивать приоритеты выдачи без необходимости пересмотра бизнес-логики. Это создаёт предпосылки для снижения нагрузки на персонал, ускорения обработки заказов и повышения конверсии, особенно в период изменения товарного предложения.

Перспективным направлением развития можно считать внедрение онлайн-обучения на стриминговых данных и расширение модели за счёт внешних признаков: ценовых трендов, пользовательских оценок, динамики конкурентов. Также может быть полезно усиление интерпретируемости модели для конечных пользователей и менеджеров, отвечающих за подбор. Рассматриваемый в статье подход способен стать основой для адаптивной рекомендательной платформы, ориентированной на быстро изменяющиеся условия рынка и устойчивое масштабируемое развитие.

Список литературы:

1. Alhijawi E., Kilani Y. The Recommender System: A Survey // International Journal of Advanced Intelligence Paradigms. – 2020. – Vol. 16, No. 1. – P. 35–53.
2. A Survey on Modern Recommendation System Based on Big Data [Электронный ресурс]. – Режим доступа: <https://arxiv.org/abs/2202.02314>
3. A Survey on Recommendation System Techniques // International Journal of Engineering Trends and Technology (IJETT). – 2023. – Vol. 71, No. 3. – P. 188–193.
4. Guolin K., Meng Q., Finley T., Wang T., Chen W., Ma W., Ye Q., Liu T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 3146–3154.
5. LightGBM documentation. LGBMRanker [Электронный ресурс] // Microsoft. – Режим доступа: <https://lightgbm.readthedocs.io/en/latest/pythonapi/lightgbm.LGBMRanker.html>