

Библиографическое описание согласно ГОСТ Р 7.0.100–2018

Мезин, Е. А. Прототип гибридной интеллектуальной системы массовой оценки рыночной стоимости / Е. А. Мезин. – Текст : непосредственный. – DOI 10.15593/2499-9873/2025.2.08 // Прикладная математика и вопросы управления / Applied Mathematics and Control Sciences. – 2025. – № 2. – С. 117–130.



**пермский
политех**

**ПРИКЛАДНАЯ МАТЕМАТИКА
И ВОПРОСЫ УПРАВЛЕНИЯ**

№ 2, 2025

<https://ered.pstu.ru/index.php/amcs>



Научная статья

DOI: 10.15593/2499-9873/2025.2.08

УДК 338.51:004



Прототип гибридной интеллектуальной системы массовой оценки рыночной стоимости

Е.А. Мезин^{1,2}

¹Пермский национальный исследовательский политехнический университет,
Пермь, Российская Федерация

²Мезекс.Информационные системы, Пермь, Российская Федерация

О СТАТЬЕ

Получена: 05 июня 2025

Одобрена: 07 июня 2025

Принята к публикации:

18 августа 2025

Финансирование

Исследование частично выполнено при финансовой поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках конкурса инновационных проектов «СТАРТ-1».

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Вклад автора

100 %.

Ключевые слова:

организационные системы, массовая оценка, поддержанные автомобили, машинное обучение, гибридные интеллектуальные системы, экспертные знания, вторичное обучение, вторичный рынок, транспортные средства, Big Data.

АННОТАЦИЯ

Предложен гибридный подход, сочетающий машинное обучение и экспертную оценку, для массовой оценки рыночной стоимости объектов, представленных на вторичном рынке. Апробация подхода выполнена на примере оценки поддержанных легковых транспортных средств. Описан разработанный прототип интеллектуальной системы, включающий модуль сбора и обработки рыночных данных, модель CatBoost, а также механизм «вторичной разметки» данных с участием профессиональных экспертов. Новизна заключается в организационной модели взаимодействия с экспертами – профессиональными участниками рынка и динамическом уточнении модели за счет оценки готовности к торгу. В результате тестирования прототипа достигнута точность на уровне MAPE $\approx 9,8\%$, что позволяет использовать систему в банковской, страховой и оценочной деятельности. Обсуждаются направления дальнейшего развития, включая NLP и CV-анализ объявлений.



Perm Polytech Style Mezin, E.A. Hybrid intelligent system prototype for mass appraisal of market value. *Applied Mathematics and Control Sciences*. 2025, no. 2, pp. 117–130. DOI: 10.15593/2499-9873/2025.2.08

MDPI and ACS Style: Mezin, E.A. Hybrid intelligent system prototype for mass appraisal of market value. *Appl. Math. Control Sci.* 2025, 2, 117–130. <https://doi.org/10.15593/2499-9873/2025.2.08>

Chicago/Turabian Style: Mezin, Evgeniy A. 2025. "Hybrid intelligent system prototype for mass appraisal of market value". *Appl. Math. Control Sci.* no. 2: 117–130. <https://doi.org/10.15593/2499-9873/2025.2.08>



APPLIED MATHEMATICS
AND CONTROL SCIENCES

№ 2, 2025

<https://ered.pstu.ru/index.php/amcs>



Article

DOI: 10.15593/2499-9873/2025.2.08

UDC 338.51 : 004



Hybrid intelligent system prototype for mass appraisal of market value

E.A. Mezin^{1,2}

¹Perm National Research Polytechnic University, Perm, Russian Federation

²Mezeks.Information Systems, Perm, Russian Federation

ARTICLE INFO

Received: 05 June 2025
Approved: 07 July 2025
Accepted for publication:
08 August 2025

Funding

The research was partially carried out with the financial support of the Foundation for Assistance to Small Innovative Enterprises within the framework of the START-1 innovation project competition.

Conflicts of Interest

The author declares no conflict of interest.

Author Contributions

100 %.

Keywords:

organizational systems, mass evaluation, used cars, machine learning, hybrid intelligent systems, expert knowledge, secondary learning, secondary market, vehicles, Big Data.

ABSTRACT

A hybrid approach combining machine learning and expert assessment is proposed for mass assessment of the market value of objects presented on the secondary market. The approach was tested using the example of assessment of used passenger vehicles. The developed prototype of an intelligent system is described, including a module for collecting and processing market data, the CatBoost model, and a mechanism for "secondary labeling" of data with the participation of professional experts. The novelty lies in the organizational model of interaction with experts - professional market participants and dynamic refinement of the model due to the assessment of readiness for bargaining. As a result of testing the prototype, an accuracy of MAPE $\approx 9.8\%$ was achieved, which allows using the system in banking, insurance and appraisal activities. Directions for further development are discussed, including NLP and CV analysis of ads.



Введение

Традиционные методы оценки основаны на трудоемком сравнительном анализе аналогов и субъективном мнении эксперта, что ограничивает масштабируемость и точность при массовом применении. Современные автоматизированные сервисы в основном ориентированы на прогнозирование цены размещения на досках объявлений, а не на фактическую цену сделок, что снижает применимость их результатов в банковской, страховой и оценочной деятельности. Ошибки в оценке могут привести к существенным финансовым потерям: от убытков при купле-продаже до недооценки залогов в банковской и лизинговой сферах.

Кроме того, нормативные документы¹, в частности, федеральные стандарты оценки², предписывают опираться на данные о состоявшихся сделках, которые при этом недоступны в открытом доступе. Это создает методологический разрыв между регуляторными требованиями и техническими возможностями.

Методы машинного обучения демонстрируют высокие результаты в прогнозировании цен, но сталкиваются с ключевым ограничением: дефицитом достоверных таргетных данных, отражающих реальную рыночную стоимость объектов-аналогов. Большинство исследователей и аналитиков обучают свои модели на ценах из объявлений, что приводит к смещению прогнозных значений и снижению их практической значимости.

Таким образом, возникает потребность в научно обоснованной системе массовой оценки, способной восполнить недостаток реальных данных и повысить точность прогнозов. В данной работе предложен гибридный подход, сочетающий анализ Big Data и участие экспертов, реализованный в виде прототипа интеллектуальной системы оценки рыночной стоимости объектов на вторичном рынке. В качестве прикладной сферы применения разработанных методов и инструментальных средств выбран российский рынок поддержанных легковых автомобилей, который как в России, так и в мире демонстрирует устойчивый рост и высокую степень изменчивости, что обостряет проблему точной и своевременной оценки их рыночной стоимости.

Исходя из этого, можно сформулировать цель, объект и предмет исследования.

Цель исследования – научное обоснование и разработка интеллектуальной информационной системы массовой оценки рыночной стоимости поддержанных автомобилей на вторичном рынке с использованием гибридных технологий, интегрирующих машинное обучение и экспертные суждения.

Объект исследования – интеллектуальная информационная система массовой оценки рыночной стоимости поддержанных автомобилей.

¹ Федеральный закон от 29.07.1998 № 135-ФЗ (ред. от 19.12.2022) «Об оценочной деятельности в Российской Федерации». В данном виде документ опубликован не был. Первоначальный текст документа опубликован в изданиях «Собрание законодательства РФ», 03.08.1998, № 31, ст. 3813, «Российская газета», 06.08.1998, № 148-149.

² Приказ Минэкономразвития России от 14.04.2022 № 200 (ред. от 30.11.2022) «Об утверждении федеральных стандартов оценки и о внесении изменений в некоторые приказы Минэкономразвития России о федеральных стандартах оценки» (вместе с «Федеральным стандартом оценки „Структура федеральных стандартов оценки и основные понятия, используемые в федеральных стандартах оценки (ФСО I)“», «Федеральным стандартом оценки „Виды стоимости (ФСО II)“», «Федеральным стандартом оценки „Процесс оценки (ФСО III)“», «Федеральным стандартом оценки „Задание на оценку (ФСО IV)“», «Федеральным стандартом оценки „Подходы и методы оценки (ФСО V)“», «Федеральным стандартом оценки „Отчетом об оценке (ФСО VI)“»). В данном виде документ опубликован не был. Первоначальный текст документа опубликован 11.05.2022 в издании «Официальный интернет-портал правовой информации»: <http://pravo.gov.ru>.

Предмет исследования – модели, методы и информационные технологии массовой оценки рыночной стоимости, основанные на совместной обработке результатов машинного обучения и экспертной информации о реальных сделках.

Теория

Для практических задач оценки (банковская, страховая, залоговая деятельность) критично получать приближение к цене фактической сделки, однако в открытых источниках доступна главным образом цена объявления, также известная как цена предложения. Это системно смещает таргет и ограничивает применимость стандартных автоматизированных подходов. В ответ на этот разрыв рассматривается гибридный способ массовой оценки, в котором прогнозная модель машинного обучения дополняется выборочной экспертной верификацией и итеративным уточнением прогнозируемых значений.

Идея состоит в том, что интеллектуальная информационная система формирует первичный прогноз на основе агрегированных рыночных данных и целенаправленно отбирает объекты с наибольшей неопределенностью для вторичной разметки у профессиональных участников рынка. Полученные экспертные суждения фиксируются как уточненные метки и используются для поточечного переобучения модели, что позволяет постепенно приближать расчетные значения стоимости к условиям реальных сделок при минимально необходимом объеме ручной разметки.

Для этого предусмотрен организационно-технический контур: интерфейсы взаимодействия с экспертами (веб-панель/бот/API), буфер уточненных меток с учетом версионирования, процедуры контроля качества и распределение задач по рейтингу экспертов, основанному на точности их разметки. Такой контур повышает устойчивость к шуму и смещениям данных, обеспечивает адаптацию к текущей рыночной конъюнктуре без остановки сервиса и обеспечивает масштабируемость на сегменты с высокой ценовой вариативностью (включая рынок подержанных транспортных средств).

В последние годы задача автоматизированной массовой оценки рыночной стоимости подержанных автомобилей активно изучается в рамках прикладных научных исследований. Существующие подходы к массовой оценке можно условно разделить на три группы:

1. Регрессионные модели – от классической линейной регрессии до алгоритмов ближайших соседей. Эти методы выявляют зависимости между характеристиками автомобиля и его ценой, однако часто страдают от переобучения и невысокой точности при сложных нелинейных связях.

Ансамблевые методы – алгоритмы, объединяющие прогнозы нескольких моделей (например, случайный лес, XGBoost, LightGBM). Они обладают высокой точностью, устойчивостью и являются одними из наиболее популярных решений в задачах массовой оценки.

2. Нейросетевые подходы – глубокие модели на основе многослойных нейросетей, способные аппроксимировать сложные зависимости. Применяются реже из-за сложности настройки, требований к данным и меньшей интерпретируемости.

Ряд публикаций [1–8] подтверждает эффективность перечисленных методов, особенно ансамблевых моделей. Однако почти все существующие исследования используют в качестве целевой переменной цену из открытых объявлений, а не цену фактической сделки. Это приводит к смещению прогнозных значений и ограничивает практическое применение таких моделей в практической деятельности.

Кроме того, часто в исследованиях и публикациях не описываются методики сбора данных, не анализируется однородность выборок и отсутствует проверка устойчивости моделей при изменении периода, региона или поставщика данных. Практическая применимость таких решений остается низкой, несмотря на хорошие метрики на ограниченных тестовых выборках.

Сравнительный анализ работ показывает, что наилучшие результаты достигаются с использованием ансамблевых методов и гибридных моделей. Так, П. Понмалар и А. Кристинал [1] рекомендуют комбинирование моделей, отмечая перспективность XGBoost и Random Forest. С.К. Сатапати с соавт. [2] продемонстрировали точность XGBoost ~92 % при $MSE = 2,64$, в то время как случайный лес показал ~90 % при ошибке 3,56.

В других работах исследовались смешанные модели: С. Хан с соавт. [3] использовали взвешенную регрессионную схему, комбинирующую RF и XGBoost, показав улучшение метрик по сравнению с одиночными моделями. Р. Гупта с соавт. [4] подтвердили высокую точность RF по сравнению с ElasticNet, SVM и деревьями решений. С. Шах с соавт. [5] подчеркивают важность качественной импутации пропусков: применение алгоритма MICE в сочетании с RF повысило точность до 94 %.

Отдельное внимание уделено градиентному бустингу: Й. Ли и соавт. [6] сравнивали LightGBM и RF, выявив преимущество LightGBM при ограниченном наборе признаков. Л. Ли и З. Йе [7] показали, что стековая модель (stacking) из GBRT, XGBoost и LightGBM дает на ~2,7 % лучшее MAE, чем каждая по отдельности. В то же время Р. Сутариа и Р. Джейн [8] пришли к противоположному выводу: простая модель RF показала лучшую точность в их эксперименте, обойдя сложные ансамбли.

Таким образом, в результате литературного обзора можно сделать следующие выводы, актуальные на текущий момент:

- ансамблевые модели (особенно градиентный бустинг) демонстрируют наилучшую точность;
- гибридные модели улучшают результат за счет комбинации преимуществ;
- качество данных, особенно полнота и достоверность, оказывает ключевое влияние;
- все исследователи используют в качестве таргета цену из объявлений, что снижает валидность результатов массовой оценки для задач, связанных с реальной стоимостью сделки.

Существует очевидный пробел между научными разработками и реальными потребностями: требуется новый научно-практический подход, который бы корректировал итоговую величину оценки и аппроксимировал ее к наиболее вероятной цене реальной сделки с конкретным экземпляром.

Сбор фактических данных реальных сделок и контекстных факторов поведения продавца и покупателя в настоящий момент времени недоступен, поэтому на существующем уровне развития науки и техники возникает идея вторичной разметки данных о предложениях товаров на открытом рынке экспертными знаниями о реальных сделках, чтобы модель училась прогнозировать именно цену продажи, а не только цену объявления. При этом предлагается построить организационную модель, где роль экспертов будут осуществлять менеджеры по продажам, цель которых заключается в том, чтобы получить контакт продавца, с которым наиболее вероятно совершить сделку, направленную на получение прибыли, что, в свою очередь, позволит снизить расходы на разметку обучающей базы данных.

Данные и методы

Для преодоления ограничений традиционных моделей, основанных исключительно на обработке данных из объявлений, в настоящей работе предлагается гибридный подход, объединяющий машинное обучение на больших массивах рыночных данных и вторичную экспертную оценку.

Ключевая идея заключается в использовании экспертного суждения как средства приближения результатов оценки к реальной цене сделки. Поскольку открытые данные не содержат информации о фактических ценах, модель обучается в два этапа:

1. Первичный этап – обучение на обширной выборке объявлений о продаже автомобилей для выявления общих закономерностей ценообразования.
2. Вторичный этап – корректировка модели на основе экспертных суждений, имитирующих цену сделки, и формирование поправочных коэффициентов (коэффициентов торга).

Экспертная разметка осуществляется специалистами по купле-продаже автомобилей, которые заинтересованы в точной оценке, поскольку результат влияет на качество входящего потока клиентов (например, при лидогенерации). Такой подход формирует естественную мотивацию экспертов давать адекватные и взвешенные оценки.

Особенность предложенного механизма – система активного экспертного обучения:

- система автоматически выбирает наиболее неопределенные объекты (по метрике энтропии, разбросу расчетных значений и т.п.);
- направляет их на экспертную оценку;
- на основе полученных значений формирует скорректированную выборку;
- повторно обучает модель, снижая смещение, связанное с ценой предложения.

При обеспечении условия, что экспертное обучение не будет требовать больших ресурсов (временных, финансовых, человеческих), система будет обладать алгоритмической масштабируемостью и гибкостью машинного обучения с точностью и интуицией профессионалов. Формируемые на основе экспертов коэффициенты торга агрегируются по кластерам (в случае оценки транспортных средств к параметрам, определяющим кластеры, относятся марка, модель, год, пробег и т.д.) и обеспечивают дополнительный уровень адаптации модели к рыночной реальности.

Итоговая цель предлагаемого подхода – построение организационно-технической системы, в которой взаимодействуют три ключевых элемента:

- 1) автоматизированный сбор данных;
- 2) модель машинного обучения;
- 3) активный контур экспертной корректировки, основанный на реальных интересах участников рынка.

В рамках этой системы независимые эксперты (менеджеры по продажам, автооценщики, «трейд-ин»-специалисты) получают для оценки случайно отобранные экземпляры автомобилей. Каждый эксперт дает свое мнение о вероятной цене реализации в текущих рыночных условиях. Эти значения используются как приближенные таргеты – аналоги фактических цен сделок.

На основе этой информации система рассчитывает коэффициенты торга – поправочные множители между ценой предложения и ожидаемой ценой продажи. Такие коэффициенты агрегируются по кластерам (например, в случае оценки транспортных средств: «KIA Rio, 2017 г.») и применяются для уточнения базовой модели. Обновление происходит итеративно, в режиме «живого» переобучения.

Ключевая особенность подхода – естественная мотивация экспертов: те, кто дают более точные оценки, получают приоритетный доступ к целевому трафику (лидогенерация от системы), а качество их работы влияет на персональный рейтинг. Это позволяет системе саморегулироваться без внешнего контроля, на основе взаимной заинтересованности всех участников.

Гибридная модель реализует субъект-объектное взаимодействие, в котором алгоритмы и люди формируют единое интеллектуальное целое, способное адаптироваться к изменяющейся рыночной среде. Подход позволяет не только повысить точность оценки, но и существенно снизить затраты на разметку данных за счет встраивания системы в бизнес-процессы профессиональных участников рынка, в рассматриваемом примере – автодилеров.

Важно подчеркнуть, что отбор экземпляров для экспертной оценки осуществлялся в автоматизированном режиме. Для эффективной работы необходимо внедрить механизм мотивации экспертов через интеграцию системы с бизнес-процессами автодилеров. Так, система может выступать лидогенератором, передавая дилерам информацию о «недооцененных» автомобилях на рынке взамен на их вклад в разметку данных.

На этапе определения исходных ценообразующих факторов для массовой оценки использовались как литературные источники [13–15], так и результаты экспертного опроса, выполненного автором настоящего исследования в рамках НИОКР [16]. В ходе исследования сформирован обобщенный список характеристик автомобиля, влияющих на его цену, и проведено анкетирование практикующих оценщиков для ранжирования этих факторов по степени их значимости. В результате выполнения НИОКР [16] ценообразующие факторы были разделены на классифицирующие и характеризующие, позже названные в работе [17] классифицирующими и истинно-ценообразующими соответственно.

Модель

Практическая значимость исследования сводится к разработке и реализации программного комплекса, реализующего научную задачу определения наиболее вероятностной цены продажи объекта оценки на вторичном рынке [18].

На рис. 1 показана общая архитектура прототипа гибридной системы массовой оценки.

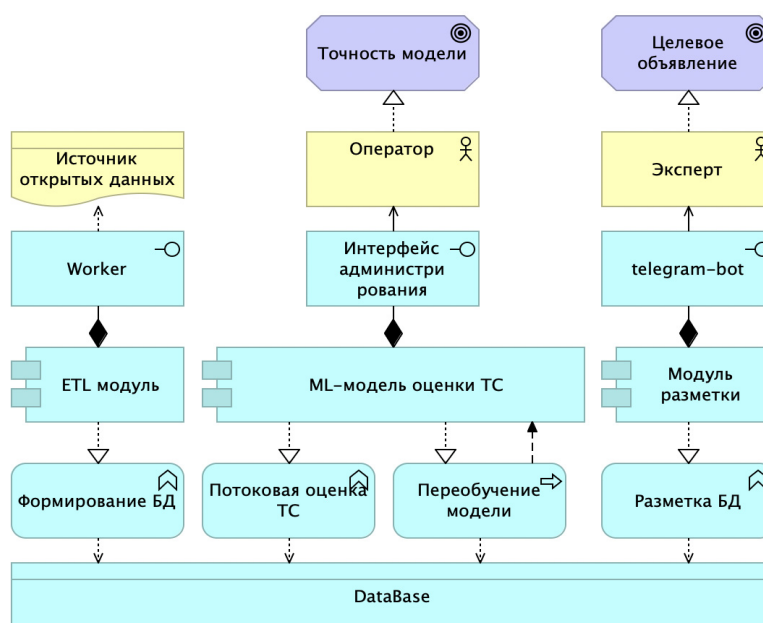


Рис. 1. Архитектура прототипа гибридной системы массовой оценки

Архитектура ориентирована на автоматическую агрегацию рыночной информации, построение модели оценки, взаимодействие с экспертами и постоянное уточнение предсказаний.

Архитектура прототипа гибридной системы массовой оценки состоит из нескольких взаимосвязанных компонентов.

Модуль сбора данных (ETL-модуль): предназначен для агрегирования информации о поддержанных автомобилях из открытых источников (веб-сайтов объявлений о продаже). Реализовано автоматизированное веб-сканирование (web-scraping), используя интерпретируемый язык Python и его библиотеки для парсинга HTML-страниц (такие как requests, BeautifulSoup, фреймворк Scrapy и др.), что соответствует лучшим практикам, описанным в литературе. Перед началом разработки были изучены различные подходы к извлечению данных: от использования headless-браузеров (Selenium) до парсинга DOM через библиотеки lxml и CSS-селекторы. Для каждого объявления сохраняются все необходимые, значимые для оценки атрибуты. Собранные данные проходят предобработку: очистку от дубликатов, нормализацию форматов (например, пробег в км, цена в рублях), приведение категориальных признаков к унифицированному виду и т.д., определена работа с пропусками и выбросами.

Все структурированные данные сохраняются в реляционной базе данных [19], реализованной в СУБД PostgreSQL. База данных [19] спроектирована с учетом предметной области: созданы связанные таблицы для марок, моделей, поколений, технических характеристик, цен и пр. Связи должны обеспечивать устранение избыточности (например, справочник марок и моделей) и удобство выборки данных по нужным критериям. Выбор в пользу PostgreSQL обусловлен его надежностью и масштабируемостью. Пополнение происходит ежедневно (средняя задержка обновления данных от актуального состояния рынка не должна превышать одного дня). База данных также хранит результаты экспертных оценок.

Модуль аналитики и машинного обучения является ядром системы, реализующим задачу оценки рыночной стоимости объекта оценки цены. Для разработки и тренировки моделей используется язык Python как наиболее популярный в сфере Data Science, обладающий богатой экосистемой библиотек. В частности, применяются: NumPy и Pandas для обработки и анализа больших табличных данных, scikit-learn и специализированные библиотеки градиентного бустинга (CatBoost), а также инструменты визуализации (Matplotlib, Seaborn) для исследования данных. На этапе исследований были опробованы различные алгоритмы машинного обучения, о которых шла речь в обзоре. По результатам сравнительного эксперимента с использованием накопленных данных был выбран метод градиентного бустинга решений, а именно алгоритм CatBoost. Данная реализация стохастического градиентного бустинга от компании «Яндекс» продемонстрировала лучшую точность и скорость сходимости по сравнению с популярным XGBoost, особенно на данных схожих задач. Существенное преимущество CatBoost – эффективная работа с категориальными признаками: алгоритм автоматически обрабатывает категориальные переменные без унитарного кодирования (One-Hot encoding), разделяя пространство решений по категориям и тем самым лучше улавливая нелинейные зависимости. Кроме того, на тестовых данных CatBoost дал высокое качество оценки даже без тщательной ручной настройки гиперпараметров, что ускорило внедрение модели. Недостатком является разве что требовательность к объему обучающей выборки, но в настоящей задаче это не критично благодаря миллионам наблюдений, собранным в базе данных. Для полноты эксперимента тестировалась также простая полносвязанная нейронная сеть (Multi-Layer Perceptron) с од-

ним выходным нейроном (рыночной стоимостью) – она тоже способна аппроксимировать зависимость, но потребовала больше времени на обучение и тщательного подбора архитектуры и параметров, в то время как CatBoost оказался быстрее в адаптации и более интерпретируемым. Поэтому финальная модель должна базироваться на CatBoost.

Перед обучением модельного алгоритма проводится подготовка данных: заполнение отсутствующих значений (используя медианы или продвинутые методы типа MICE), удаление или коррекция выбросов (например, нереалистично малый пробег для старого авто – такие аномалии отсеиваются на основе статистических правил в 3 сигмы), масштабирование числовых признаков по необходимости. Категориальные признаки, не поддерживаемые напрямую моделью, кодируются в зависимости от алгоритма: для CatBoost значительная часть категорий обрабатывается внутренне, а редкие значения сводятся к категории «другое». Обучающая выборка формируется за выбранный период (например, последние 3 месяца рынка, чтобы модель актуально отражала текущие ценовые тенденции). Части выборки: основная масса данных с ценами объявлений и сравнительно небольшое число записей с ценами, скорректированными экспертами.

Модуль экспертной коррекции. Сервис отвечает за взаимодействие с экспертами и обработку поступающих от них оценок. В прототипе реализован простой телеграмм бот для экспертов, где им в личном кабинете предлагается набор ссылок на объявления, подлежащих оценке (формируемый системой автоматически). Эксперт просматривает описание автомобиля (марка, модель, год, пробег, фото и т.д.) и вводит свою оценку – сумму, за которую, по его мнению, данный автомобиль реально может быть продан в течение месяца. Система аккумулирует ответы экспертов и строит суждение о коэффициенте торга, справедливом для всего рынка в целом и для локальных отклонений в рамках отдельных когорт «марка – модель – поколение» автомобиля (рис. 2).

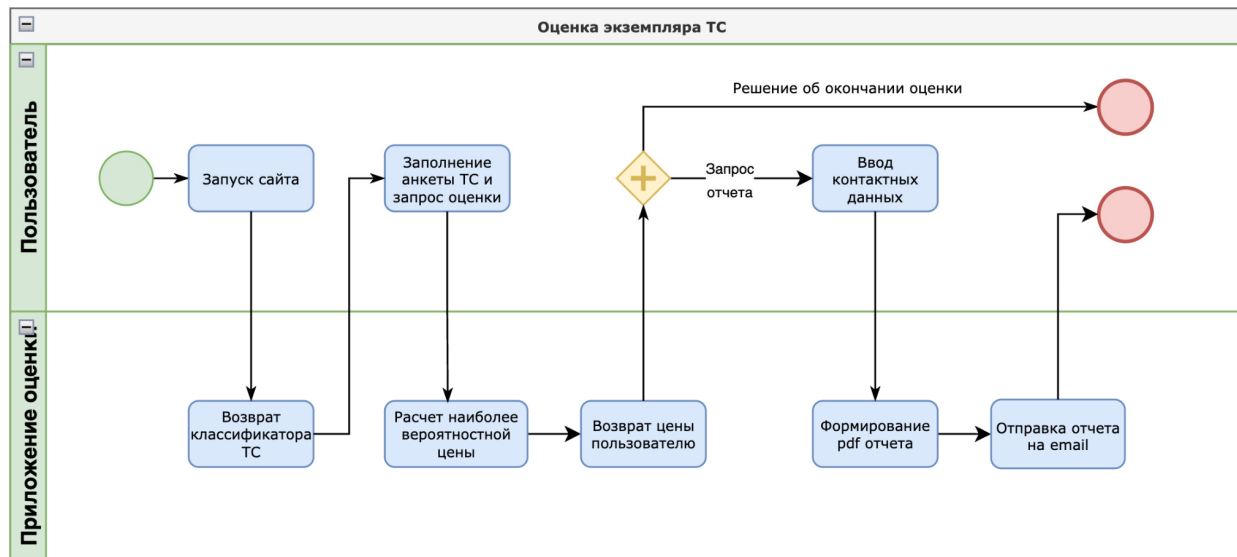


Рис. 2. Общая модель взаимодействия пользователя и приложения

Полученные результаты

1. Качество прогнозирования. Для оценки точности разработанной модели использовалась метрика средней абсолютной процентной ошибки (MAPE) между предсказанной ценой и фактической (в нашем случае – либо известной ценой сделки, либо экспертно

скорректированной ценой объявления). Согласно техническому заданию, требовалось достичь MAPE не более 10 %.

По итогам тестирования гибридная модель показала ошибку $\approx 9,8$ %, что соответствует целевым требованиям и подтверждает пригодность модели к практическому применению в задачах оценки залоговой стоимости автомобилей, «трейд-ин» и страховании.

2. Эффект вторичного обучения. Особый вклад в повышение точности дал механизм вторичной коррекции на основе экспертных данных. Даже при использовании относительно небольшой выборки (~ 4000 оценок) модель существенно улучшила качество предсказаний в сегментах с систематическим отклонением между ценой объявления и ценой сделки.

В частности, модель научилась дисконтировать завышенные предложения и корректно предсказывать цену, приближенную к ожидаемой цене реализации. Это особенно важно для практик, связанных с рисками недооценки или переоценки залогов.

3. Пользовательская апробация. Для оценки практической значимости прототип был представлен пилотной группе пользователей: сотрудникам банков, автодилерам и независимым оценщикам. В рамках тестирования участникам предлагалось ввести данные автомобилей и оценить адекватность полученных результатов.

Большинство респондентов отметили:

- высокую скорость обработки;
- релевантность прогнозируемой стоимости;
- удобство интерфейса.

Среди предложений по улучшению: добавить сравнение с аналогами (подобные объявления), ярче учитывать региональные особенности, предоставить детализированный отчет для клиента.

Эти замечания были зафиксированы и будут учтены при доработке промышленной версии гибридной интеллектуальной системы массовой оценки рыночной стоимости.

Заключение

В работе предложен и реализован подход к массовой оценке рыночной стоимости подержанных автомобилей на вторичном рынке, основанный на гибридной интеллектуальной системе. Архитектура прототипа [18] включает модули сбора данных, машинного обучения и экспертной коррекции, направленной на реализацию механизма вторичного обучения, а также базу данных [19].

В условиях дефицита информации о реальных ценах сделок предложено использовать экспертные оценки как приближенные целевые значения, что позволило повысить точность массовой оценки. Экспериментально достигнута средняя абсолютная процентная ошибка $\approx 9,8$ %, что соответствует требованиям практического применения в сфере банков, автодилеров, страхования и оценки.

Прототип успешно апробирован с участием профессиональных пользователей, которые отметили ее релевантность, быстродействие и удобство. Это подтверждает жизнеспособность подхода и потенциал его масштабирования [16].

Планируемые направления дальнейших исследований включают:

- интеграцию методов анализа изображений (CV) для оценки состояния автомобилей по фотографиям;
- применение алгоритмов обработки естественного языка (NLP) для анализа текстов объявлений;

- расширение географии и динамическое обновление экспертной базы;
- разработку интерфейсов API для внешней интеграции;
- создание метрик качества экспертов.

Предложенная система формирует предпосылки для развития новой парадигмы в оценочной деятельности – на стыке машинного обучения, организационных механизмов и коллективного экспертного интеллекта. Это открывает перспективу создания надежного, масштабируемого и гибкого инструмента оценки даже в условиях ограниченности данных.

Кроме того, созданная система обладает высоким потенциалом масштабируемости как на другие категории транспортных средств (коммерческий транспорт, мотоциклы), так и на смежные рынки, где проблема отсутствия данных о реальных сделках также критична (недвижимость, оборудование, техника). Встраивание механизма активного обучения в бизнес-процессы дает основание полагать, что предложенный подход может лечь в основу нового поколения оценочных платформ, ориентированных на данные и экспертную обратную связь.

Список литературы

1. Ponmalar, P. Review on the Pre-owned Car Price Determination using ML Approaches / P. Ponmalar, A. Christinal // Proceedings of 2022 International Conference on Artificial Intelligence and Smart Systems (ICAISS). – IEEE, 2022. – P. 274–278. DOI: 10.1109/ICAISS55157.2022.10010958
2. Satapathy, S.K. An Automated Car Price Prediction System Using Effective ML Techniques / S.K. Satapathy, R. Vala, S. Virpariya // Proceedings of 2022 International Conference on Inventive Computation and Information Technologies (CISES). – IEEE, 2022. – P. 402–408. DOI: 10.1109/CISES54857.2022.9844350
3. Han, S., Second-hand Car Price Prediction Based on a Mixed-Weighted Regression Model / S. Han, J. Qu, J. Song, Z. Liu // Proceedings of 2022 7th International Conference on Big Data Analysis (ICBDA). – IEEE, 2022. – P. 90–95. DOI: 10.1109/ICBDA55095.2022.9760371
4. Automobile Price Prediction using Regression Models / R. Gupta, A. Sharma, V. Anand, S. Gupta // Proceedings of 2022 International Conference on Information and Communication Technology (ICICT). – IEEE, 2022. – P. 410–416. DOI: 10.1109/ICICT54344.2022.9850657
5. Shah, S. Imputing Missing Values for Dataset of Used Cars / S. Shah, M. Telrandhe, P. Waghmode, S. Ghane // Proceedings of 2022 International Conference on Asian Convergence Technology (ASIANCON). – IEEE, 2022. – P. 1–5. DOI: 10.1109/ASIANCON55314.2022.9908600
6. Li, Y. Research on Used Car Price Prediction Based on RF and LightGBM / Y. Li, Y. Li, Y. Liu // Proceedings of 2022 International Conference on Data Science and Computer Application (ICDSCA). – IEEE, 2022. – P. 539–543. DOI: 10.1109/ICDSCA56264.2022.9988116
7. Li, L. Research on Used Car Price Prediction Based on Stacking Model Fusion / L. Li, Z. Ye // Proceedings of 2022 International Conference on Information Network and Communication (ICINC). – IEEE, 2022. – P. 86–90. DOI: 10.1109/ICINC58035.2022.00025
8. Sutaria, R. Auto-Price Forecast: Analysis of Car Value Trends / R. Sutaria, R. Jain // Proceedings of 2023 International Conference on Emerging Trends in Engineering and Technology (INCET). – IEEE, 2023. – P. 1–6. DOI: 10.1109/INCET57972.2023.10170263
9. Ifthikar, A. Valuation of Used Vehicles: A Computational Intelligence Approach / A. Ifthikar, K. Vidanage // Proceedings of 8th International Conference on Intelligent Systems, Modelling and Simulation (ISMS). – IEEE, 2018. – P. 7–10. DOI: 10.1109/ISMS.2018.00011

10. Narayana, C.V. Machine Learning Techniques to Predict the Price of Used Cars / C.V. Narayana, C.L. Likitha, S. Bademiya, K. Kusumanjali // Proceedings of 2021 International Conference on Emerging Smart Computing and Informatics (ICESC). – IEEE, 2021. – P. 1680–1687. DOI: 10.1109/ICESC51422.2021.9532845
11. Чепыгов, Е.А. Разработка интеллектуальной информационной системы для онлайн-оценки стоимости транспортного средства / Е.А. Чепыгов // Информационные технологии и математическое моделирование высокотехнологичных систем: материалы Всероссийской конф. с междунар. участием, Москва, 13–17 апреля 2020 года. – М.: РУДН, 2020. – С. 218–222.
12. Дубровский, И.А. Автоматический расчет статистических параметров выборки / И.А. Дубровский // Вестник аритмологии. – 2015. – № 80. – С. 63–66.
13. Субботин, А.В. Использование искусственных нейронных сетей в задаче оценки стоимости автомобиля марки «KIA» / А.В. Субботин, А.А. Рукавишников, П.В. Корчемный // Концепция «общества знаний» в современной науке: сб. ст. Междунар. науч.-практ. конф. – Стерлитамак: Агентство междунар. исследований, 2019. – С. 193–195. – EDN SFCAGA.
14. Трифонов, Н.Ю. Оценка стоимости автомобильных транспортных средств / Н.Ю. Трифонов, Б.А. Романов // Автотранспортное предприятие. – 1998. – № 4 (16). – С. 40–43. – EDN VENDFP.
15. Хитяев, Е.А. Статистический анализ автомобильного рынка России / Е.А. Хитяев // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем: материалы Всероссийской конф. с междунар. участием. – М.: РУДН, 2021. – С. 492–496. – EDN KDWVJH.
16. Разработка и тестирование прототипа системы оценки стоимости транспортных средств на основе анализа Big Data с учетом факторов нестабильности рынков для автоматизации деятельности финансовых институтов: отчет о НИОКР (заключительный) / В.Ф. Гараев, А.А. Колчанов, Д.В. Белов, Е.А. Мезин, Е. А. Денисов; Мезекс.Информационные системы; Фонд содействия развитию малых форм предприятий в научно-технической сфере. Рег. № ЕГИСУ НИОКТР 223101200012-2. – Пермь: Мезекс.Информационные системы. – 157 с.
17. Алексеева И.Е. Декомпозиция ценообразующих факторов на основе статистического анализа экспертных оценок их значимости / И.Е. Алексеева, Е.А. Мезин, А.О. Алексеев // Вестник Южно-Уральского государственного университета. Серия «Компьютерные технологии, управление, радиоэлектроника». – 2025. – Т. 25, № 3. – С. 53–63.
18. Свидетельство о государственной регистрации программы для ЭВМ № 2023663596 Российская Федерация. Приложение оценки стоимости транспортных средств: № 2023662797 : заявл. 21.06.2023 : опубл. 26.06.2023 / Е.А. Мезин ; заявитель ООО «Мезекс. Информационные системы». – EDN ZLRFZ.
19. Свидетельство о государственной регистрации базы данных № 2023621200 Российская Федерация. База данных оценки стоимости транспортных средств: № 2023620927 : заявл. 03.04.2023 : опубл. 13.04.2023 / Е.А. Мезин ; заявитель ООО «Мезекс. Информационные системы». – EDN UBLNFI.

References

1. Ponmalar P., Christinal A. Review on the Pre-owned Car Price Determination using ML Approaches. *Proceedings of 2022 International Conference on Artificial Intelligence and Smart Systems (ICAISS)*, IEEE, 2022, pp. 274–278. DOI: 10.1109/ICAISS55157.2022.10010958.
2. Satapathy S.K., Vala R., Virpariya S. An Automated Car Price Prediction System Using Effective ML Techniques. *Proceedings of 2022 International Conference on Inventive Computation and Information Technologies (CISES)*, IEEE, 2022, pp. 402–408. DOI: 10.1109/CISES54857.2022.9844350.
3. Han S., Qu J., Song J., Liu Z. Second-hand Car Price Prediction Based on a Mixed-Weighted Regression Model. *Proceedings of 2022 7th International Conference on Big Data Analysis (ICBDA)*, IEEE, 2022, pp. 90–95. DOI: 10.1109/ICBDA55095.2022.9760371.
4. Gupta R., Sharma A., Anand V., Gupta S. Automobile Price Prediction using Regression Models. *Proceedings of 2022 International Conference on Information and Communication Technology (ICICT)*, IEEE, 2022, pp. 410–416. DOI: 10.1109/ICICT54344.2022.9850657.
5. Shah S., Telrandhe M., Waghmode P., Ghane S. Imputing Missing Values for Dataset of Used Cars. *Proceedings of 2022 International Conference on Asian Convergence Technology (ASIANCON)*, IEEE, 2022, pp. 1–5. DOI: 10.1109/ASIANCON55314.2022.9908600.
6. Li Y., Li Y., Liu Y. Research on Used Car Price Prediction Based on RF and LightGBM. *Proceedings of 2022 International Conference on Data Science and Computer Application (ICDSCA)*, IEEE, 2022, pp. 539–543. DOI: 10.1109/ICDSCA56264.2022.9988116.
7. Li L., Ye Z. Research on Used Car Price Prediction Based on Stacking Model Fusion. *Proceedings of 2022 International Conference on Information Network and Communication (ICINC)*, IEEE, 2022, pp. 86–90. DOI: 10.1109/ICINC58035.2022.00025.
8. Sutaria R., Jain R. Auto-Price Forecast: Analysis of Car Value Trends. *Proceedings of 2023 International Conference on Emerging Trends in Engineering and Technology (INCET)*, IEEE, 2023, pp. 1–6. DOI: 10.1109/INCET57972.2023.10170263.
9. Ifthikar A., Vidanage K. Valuation of Used Vehicles: A Computational Intelligence Approach. *Proceedings of 8th International Conference on Intelligent Systems, Modelling and Simulation (ISMS)*, IEEE, 2018, pp. 7–10. DOI: 10.1109/ISMS.2018.00011.
10. Narayana C.V., Likitha C.L., Bademiya S., Kusumanjali K. Machine Learning Techniques to Predict the Price of Used Cars. *Proceedings of 2021 International Conference on Emerging Smart Computing and Informatics (ICESC)*, IEEE, 2021, pp. 1680–1687. DOI: 10.1109/ICESC51422.2021.9532845.
11. Chepygov E.A. Development of an Intelligent Information System for Online Vehicle Valuation. *Information Technologies and Mathematical Modeling of High-Tech Systems*, Moscow, April 13–17, 2020. Moscow: RUDN University, 2020, pp. 218–222.
12. Dubrovskii I.A. Automatic Calculation of Statistical Parameters of a Sample. *Bulletin of Arithmology*, 2015, no. 80, pp. 63–66.
13. Subbotin A.V., Rukavishnikov A.A., Korchemnyi P.V. Using of Artificial Neural Networks in the Task of Estimating the Cost of a KIA Car. *The Concept of "Knowledge Society" in Modern Science: Coll. of art. of the Intern. Sci.-Pract. Conf.* Sterlitamak: Agentstvo mezhdunarodnykh issledovaniy, 2019, pp. 193–195. EDN SFCAGA.
14. Trifonov N.Yu., Romanov B.A. Valuation of Motor Vehicles. *Motor Transport Enterprise*, 1998, no. 4 (16), pp. 40–43. EDN VENDFP.
15. Khityaev E.A. Statistical Analysis of the Russian Automobile Market. *Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems: Proc. of*

the All-Russian Conf. with international participation. Moscow: RUDN University, 2021, pp. 492–496. EDN KDWVJH.

16. Garaev V.F., Kolchanov A.A., Belov D.V., Mezin E.A., Denisov E.A. Razrabotka i testirovanie prototipa sistemy otsenki stoimosti transportnykh sredstv na osnove analiza Big Data s uchetom faktorov nestabilnosti rynkov dlya avtomatizatsii deyatel'nosti finansovykh institutov [Development and Testing of a Prototype System for Vehicle Valuation Based on Big Data Analysis Considering Market Instability Factors for Automation of Financial Institutions]: Final R&D Report. Perm, Mezeks.Informatsionnye sistemy, 2022, 157 p.

17. Alekseeva I.E., Mezin E.A., Alekseev A.O. Decomposition of Price-Forming Factors Based on Statistical Analysis of Expert Assessments of Their Significance. *Bulletin of the South Ural State University. Series "Computer Technologies, Control, Radio Electronics"*, 2025, vol. 25, no. 3, pp. 53–63.

18. Mezin E.A. Prilozhenie otsenki stoimosti transportnykh sredstv [Vehicle Valuation Application] Certificate of the Computer Program RU 2023662797 26 June 2023, Russian Federation.

19. Mezin E.A. Baza dannykh otsenki stoimosti transportnykh sredstv [Vehicle Valuation Database: No. 2023620927] Certificate of the Database RU 2023620927 from 13 April 2023, Russian Federation.