

ИНФОРМАТИКА И ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ/INFORMATICS AND INFORMATION PROCESSES

DOI: <https://doi.org/10.60797/itech.2026.9.2>

ОЦЕНКА СТОИМОСТИ АВТОМОБИЛЕЙ С ПРИМЕНЕНИЕМ МЕТОДОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Научная статья

Чижов А.С.^{1,*}

¹ Ивановский государственный химико-технологический университет, Иваново, Российская Федерация

* Корреспондирующий автор (alexchizh02[at]yandex.ru)

Аннотация

Если сравнивать рынок подержанных и новых автомобилей, то очевидно, что первый намного больше. В последние годы данный рынок достаточно непредсказуем из-за чего становится очень рисковым. Для помощи в принятии верного решения как покупателю, так и продавцу помогут несколько моделей машинного обучения. Целью работы является определение основных факторов, влияющих на стоимость авто, а также оценка их влияния с помощью алгоритмов искусственного интеллекта. Результат проведенной работы показал, что наиболее влияющими факторами на цену автомобиля являются пробег, возраст, марка, техническое состояние и экономическая ситуация. Полученные данные могут быть полезными для понимания реальной рыночной стоимости подержанного авто.

Ключевые слова: машинное обучение, анализ данных, предсказание цен, автомобили, интеллектуальный анализ.

EVALUATION OF THE VEHICLE VALUE USING ARTIFICIAL INTELLIGENCE METHODS

Research article

Chizhov A.S.^{1,*}

¹ Ivanovo State University of Chemical Technology, Ivanovo, Russian Federation

* Corresponding author (alexchizh02[at]yandex.ru)

Abstract

When comparing the used and new car markets, it is clear that the former is much larger. In recent years, this market has been quite unpredictable, which makes it very risky. Several machine learning models can help both buyers and sellers make the right decision. The aim of this work is to identify the main factors affecting the cost of a car and to evaluate their impact using artificial intelligence algorithms. The results of the study showed that the most influential factors affecting the price of a car are mileage, age, brand, technical condition, and economic situation. The obtained data can be useful for understanding the real market value of a used car.

Keywords: machine learning, data analysis, price prediction, cars, intelligent analysis.

Введение

Актуальность данной работы связана с тем, что цены на авто в России на протяжении последних пяти лет сильно растут. На это влияет множество факторов, которые каждый год добавляются. Последними, с которыми пришлось столкнуться, стали: коронавирус, начало СВО, упадок великих авто брендов, множество ограничений различных стран на продажу иностранных машин. Модели машинного обучения могут анализировать различные характеристики авто и большие объемы данных. Передовые алгоритмы дают возможность высокого точного прогнозирования. Это гораздо эффективнее таких традиционных методов, как однофакторная и многомерная линейная регрессия, которые имеют некоторые минусы. Во-первых, они работают только с линейными зависимостями между переменными. Во-вторых, если данные имеют выбросы или экстремальные значения, прогноз сильно искажается. Применение автоматизированного процесса предсказания цен на автомобили поможет повысить точность прогнозов.

Цель работы: выявить ключевые факторы, влияющие на стоимость авто с применением методов машинного обучения.

Задачи, поставленные перед этой работой:

- изучение различных методов машинного обучения;
- выявление тенденций в оценке автомобилей с использованием искусственного интеллекта (ИИ) и больших данных;
- проанализировать и понять, какие факторы больше всего влияют на цену автомобилей.

Основные результаты

Машинное обучение (Machine Learning, ML) — это совокупность методов искусственного интеллекта, с помощью которых можно создавать самообучающиеся компьютерные системы (в частности, нейросети). Для таких систем разработчики не прописывают конкретные алгоритмы решения задач, а предоставляют подготовленные данные и описывают критерии успешного решения, по которым учатся нейросети [1].

В традиционном программировании для решения некой задачи разработчику необходимо определить алгоритм и «объяснить» его компьютеру с помощью кода на языке программирования. У специалистов по машинному обучению другой подход: они пытаются научить машину «думать», подобно человеку, и самостоятельно находить решения.

Рассмотрим различные методы машинного обучения, которые помогут предсказать цены на подержанные автомобили.

Первый и самый простой метод — это линейная регрессия. Её суть заключается в нахождении наилучшей прямой, которая описывает линейную связь между целевой переменной и множеством объясняющих факторов. В нашем случае объясняющими (независимыми) переменными выступают марка, возраст и пробег автомобиля, а целевой (зависимой) — его цена. Этот метод позволяет выявить общие тенденции, но даёт неточные результаты, если взаимосвязи между факторами имеют сложный, нелинейный характер [2].

Более точный прогноз можно получить, применяя деревья решений и случайный лес. Дерево решений представляет данные в мелких группах, после делает прогноз. Это выглядит как цепочка вопросов, где каждый вопрос делит данные на две части, точнее на промежуточные узлы, которые формируются на основе бинарных условий (да/нет) [3]. Пример по теме статьи:

Автомобиль старше 5 лет?

Да — следовательно, цена, скорее всего, ниже.

Нет — значит, идем дальше.

Случайный лес выглядит следующим образом. Берется много деревьев решений, работающих вместе, где каждое делает свое предсказание. После полученные результаты усредняются. Данная процедура позволяет получить более точный результат. При данных методах алгоритмы учитывают сложные нелинейные зависимости, но при этом чувствительны к шуму данных.

Самый передовой и умный способ улучшить предсказания — это градиентный бустинг. Его суть заключается в объединении множества слабых моделей, которые каждый раз повышают свою точность, опираясь на совершенные ранее ошибки. В результате получается мощная модель, которая может делать очень точные прогнозы [4].

Выделим три его главных плюса:

1. Улавливает сложные закономерности.
2. Исправляет в каждой новой модели предыдущие ошибки.
3. Может работать с разными данными и задачами.

XGBoost, LightGBM, CatBoost — популярные алгоритмы градиентного бустинга, которые показали отличные результаты в быстродействии и эффективности [5].

Еще один интересный метод — это кластеризация. В нем данные делятся на группы так, чтобы похожие объекты были рядом. Это позволяет группировать автомобили по схожим характеристикам и выявлять закономерности в их стоимости. Самый простой и популярный метод — К-средних. Он берет К случайных центров, после чего все объекты распределяют в группы и начинают постепенно уточнять границы кластеров [6].

Также могут использоваться нейросетевые методы, к примеру MLP и CNN. Тут надо понимать, что нейросети — это алгоритмы, которые работают по принципу человеческого мозга. Они делают анализ данных, затем занимаются поиском закономерностей, в результате чего можно получить прогноз. MLP работает с числами. Он берет пробег, год выпуска, количество владельцев и так далее, после чего делает расчеты и выдает итоговую цену. CNN анализирует картинки и находит важные детали. Он способен заметить царапины, повреждения и износ салона [7].

Внешние факторы, такие как цена топлива, сезонность, рыночные колебания, также очень важны для предсказания цен на автомобили. Для этого можно использовать методы временных рядов. ARIMA может, как пожилой эксперт, оценить как менялись цены в прошлом, то есть выстроить повторяющиеся закономерности и на основе этого предсказать будущее [8]. LSTM — это нейросеть способная запомнить множество разных факторов из прошлого и сопоставить их между собой для предсказания будущего [9].

Последним в изучении машинного обучения будет их интерпретируемость. Обычно нейросети и сложные алгоритмы не объясняют, как они приходят к своему выводу, что ставит под сомнение результат. Для этого есть методы объяснения решений моделей. SHAP делает разбор каждого фактора, о том какой финансовый вклад был вложен. К примеру, малый пробег +200 000 руб., старый год -100 000 руб. LIME объясняет отдельные предсказания модели, показывая, какие факторы повлияли на конкретный результат [10].

Теперь важно поговорить о будущей оценки автомобилей: роль искусственного интеллекта и больших данных.

Современные технологии в последние десятилетия развиваются с очень быстрой скоростью, поэтому стоит выделить основные тенденции оценки стоимости автомобилей.

Тенденции в оценке автомобиля с использованием ИИ:

1. Алгоритмы машинного обучения, которые описаны выше, такие как линейная регрессия и деревья решений, уже предсказывают цены на автомобили, однако если факторов становится слишком много, то начинают появляться ошибки. Решением этой проблемы и дальнейшим развитием являются глубокие нейросети. Они могут учитывать больше факторов и выстраивать сложные взаимосвязи между ними. К примеру, рекуррентные нейронные сети (RNN) и трансформеры могут анализировать изменения цен во времени, а также как они будут меняться в будущем. Про сверточные нейросети (CNN) было сказано выше.

2. Полная интеграция Интернета вещей (IoT) в оценку автомобилей. С каждым годом машины все больше оснащаются различными датчиками для сбора информации, такими как GPS, расходы различных жидкостей и так далее. Также автопроизводители постоянно их улучшают для уменьшения риска программного взлома и сброса показаний. Эти данные позволяют нейросетям оценить не только такие факторы, как пробег, но и качество эксплуатации автомобилей на каждом промежутке жизни. Покупатель, благодаря этому, сможет узнать, как ездили на машине (быстро или спокойно), часто ли меняли расходники и так далее [11].

Применение больших данных в прогнозировании стоимости автомобилей:

1. Раньше прогнозирование цен автомобиля исходило из исторических данных. Сейчас есть облачные платформы, которые хранят всю информацию о ситуации на рынке, что позволяет машинному обучению и нейросетям выдавать точные прогнозы.

2. Применение блокчейн-технологий с автомобилем. Это одна из перспективных областей в мире автомобилей, так как покупатель и технологии, описанные выше, смогут получить данные недоступные для подделки, в которых будут храниться стоимости при продаже, покупки запчастей, аварии, пробег

Автоматизация на рынке оценки автомобилей:

1. Специалисты по оценке автомобилей будут постепенно уходить в прошлое, так как все вышеописанные технологии вместе смогут предоставлять быстрый и качественный вариант и скорее всего за более низкую цену. Также важно учитывать вопрос простоты взаимодействия. Последнее время люди предпочитают хорошо настроенную онлайн платформу, нежели общение с другим человеком.

2. Уровень персонализации. Сейчас каждый сайт пытается собрать информацию о человеке, чтобы предоставить ему интересный товар и соответствующую рекламу партнера. Платформа, на которой человек оставляет свой цифровой след, намного лучше может понять потребности и пожелания, если сравнивать ее со специалистом, который потребует множество вводных данных для начала, а после может предлагать неинтересные варианты.

Будущее оценки автомобиля уже связано с ИИ и анализом больших данных. Дальше нас ждет продолжение этого развития, которое позволит делать прогнозы более точными, прозрачными, удобными для всех участников рынка.

Последняя поставленная задача — это понять, какие факторы больше всего влияют на цену автомобиля. Проанализировав различные популярные источники и результаты ИИ, анализ показал следующий топ из пяти мест [12]:

1. В возраст автомобилия. Это самый главный фактор при выборе машины, так как невозможно восстановить старую машину до состояния новой. При этом это не рентабельно: придется заплатить стоимость, равную новой машине или даже больше, но устаревший дизайн и старые технологии не обновятся.

2. Пробег. Проехав много километров, даже новая машина за короткий срок будет в изношенном состоянии, так как автопроизводители закладывают определенный жизненный цикл в такие важные части, как мотор с коробкой и кузов.

3. Марка и модель. Каждый авто бренд имеет свою историю и заслуги, что сильно влияет на стоимость как новой машины, так и подержанной. Некоторые модели становятся легендами своего времени, поэтому даже в массовом сегменте можно встретить авто, которые стоят дороже своих конкурентов и при этом пользуются, несмотря на это, большой популярностью.

4. Техническое состояние. Даже легендарная модель, побывавшая в серьезном ДТП и после восстановленная, не может стоить дорого, так как восстановить в заводской варианте невозможно, что сказывается на безопасности водителя при следующем ДТП. Не только важно состояние кузова, но и мотор с подвеской, так как двигатель авто в некоторых случаях может стоить как половина или вся машина целиком.

5. Рынок и экономические факторы. Последние годы мы можем прочувствовать этот фактор на себе, так как на нестабильность цен на автомобили оказывают колебания валют, налоговая и таможенная политика, кредитование.

Практическое применение интеллектуального анализа в автооценке

Получение прогнозируемой цены не является итогом - еще важно преобразовать полученные данные в инсайт. Сделанный анализ дает возможность выделить несколько ключевых направлений применения этих данных:

1. Получение прогнозируемой цены на автомобиль с помощью автоматизированных алгоритмов, позволяет покупателям определять реальную цену авто и избегать переплат.

2. Удобная оптимизация ценообразования для дилеров. Автосалоны смогут определить конкурентоспособные цены при использовании машинного обучения.

3. Дилеры и страховые компании смогут предсказать остаточную стоимость транспортного средства через несколько лет. Это позволит им точно рассчитать лизинговые платежи и страховые взносы.

4. Постоянное изменение внешних факторов, таких как изменение законодательства, социальные волнения, напрямую влияют на стоимость авто. Проанализировав эти факторы современными технологиями, можно адаптировать стратегии продаж.

Современные методы интеллектуального анализа данных могут не просто предсказать стоимость автомобиля, но и позволяют принимать обоснованные решения на основе реальных рыночных закономерностей.

Заключение

Использование методов интеллектуального анализа данных помогают автоматизировать, а также повысить точность оценки стоимости авто. В данной работе были изучены различные методы машинного обучения, выявлены тенденции в оценке автомобилей с использованием ИИ и больших данных и проанализированы факторы, которые больше всего влияют на цену автомобилей.

Результаты исследования могут помочь разным субъектам авторынка. Также важно учитывать, что машинное обучение не только говорит точную прогнозную цену, но и способствует автоматизированной системе оценивания, делает рынок более прозрачным и предсказуемым.

Дальнейшие исследования можно развивать в сторону более сложных факторов — к примеру, результаты автоспорта, тенденции в производстве авто и развитие задающих факторов стиля авто.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы / References

1. Как устроено машинное обучение: задачи, алгоритмы и виды Machine Learning // Skillbox. — 2023. — URL: <https://skillbox.ru/media/code/kak-ustroeno-mashinnoe-obuchenie-zadachi-algoritmy-i-vidy-machine-learning/>. (дата обращения: 01.02.25)
2. Montgomery D.S. Introduction to Linear Regression Analysis / D.S. Montgomery, E.A. Pek, G.G. Wining. — Moscow: John Wiley & Sons Limited, 2021. — 672 p.
3. Бондарев А.В. Применение алгоритмов машинного обучения в прогнозе результата пиролитического анализа / А.В. Бондарев, Е.В. Козлова // ПолиПРЕСС. — 2021. — №3. — С. 333–336. — URL: https://www.elibrary.ru/item.asp?utm_source=chatgpt.com&id=47167532 (дата обращения: 02.05.25).
4. Султанов Б.Р. Использование метода экстремального градиентного бустинга в прогнозировании консолидированного бюджета региона / Б.Р. Султанов, Д.Р. Султангареев // Экономика и управление: научно-практический журнал. — 2021. — № 5 (161). — С. 134–139. — URL: https://ekam-journal.com/images/6-2021/5-2021/Sultanov_Sultangareev.pdf?utm_source=chatgpt.com (дата обращения: 02.02.25). — DOI: 10.34773/EU.2021.5.23
5. Захаренко Е. Градиентный бустинг. Реализация с нуля на Python и разбор особенностей его модификаций (XGBoost, CatBoost, LightGBM) / Е. Захаренко // Хабр. — 2024. — URL: <https://habr.com/ru/articles/799725/> (дата обращения: 02.02.25)
6. Баданина Н.В. Кластеризация водителей по степени опасности вождения с использованием алгоритмов машинного обучения / Н.В. Баданина, В.А. Судаков // Моделирование и анализ данных. — 2022. — Т. 12. — № 1. — С. 5–15. — URL: https://psyjournals.ru/journals/mda/archive/2022_n1/Badanina_Sudakov (дата обращения: 01.02.25). — DOI: 10.17759/mda.2022120101
7. Иванова А.А. Нейросетевые технологии в системе оценки ресурсного потенциала аграрного сектора АПК / А.А. Иванова, О.Ю. Смысlova // Экономика, предпринимательство и право. — 2024. — Т. 14. — № 10. — С. 5771–5786. — URL: <https://1economic.ru/lib/121808> (дата обращения: 02.02.25). — DOI: 10.18334/epp.14.10.121808
8. Бозжигитов Б.Е. Прогнозирование временных рядов с использованием модели ARIMA / Б.Е. Бозжигитов, Ж.Ж. Кожамкулова // Основы и тенденции в исследовательской деятельности. — 2024. — № 8.
9. Ал жеев А.В. Сравнительный анализ прогнозных моделей ARIMA и LSTM на примере акций российских компаний / А.В. Ал жеев, Р.А. Кочкаров // Финансы: теория и практика. — 2020. — Т. 24. — № 1. — С. 14–23. — URL: <https://financetp.fa.ru/jour/article/view/950> (дата обращения: 03.02.25). — DOI: 10.26794/2587-5671-2020-24-1-14-23
10. Седухин О. Интерпретация моделей и диагностика сдвига данных: LIME, SHAP и Shapley Flow / О. Седухин // Хабр. — 2022. — URL: <https://habr.com/ru/companies/ods/articles/599573/> (дата обращения: 31.01.25)
11. Маторин С.И. Информационно-аналитическая технология будущего – Интернет вещей (IoT) / С.И. Маторин, С.В. Гуль // Научный результат. Информационные технологии. — 2023. — Т. 8. — № 4. — С. 70–78. — URL: <https://trrinformation.ru/journal/article/view/2518-1414-2023-8-4-0-6> (дата обращения: 03.02.25). — DOI: 10.18413/2518-1414-2023-8-4-0-6
12. Что влияет на стоимость автомобиля // АвтоКод. — 2021. — URL: <https://avtocod.ru/chto-vliyaet-na-stoimost-avtomobiliya>. (дата обращения: 05.02.25)

Список литературы на английском языке / References in English

1. Kak ustroeno mashinnoe obuchenie: zadachi, algoritmy i vidy' Machine Learning [How Machine Learning Works: Tasks, Algorithms, and Types] // Skillbox. — 2023. — URL: <https://skillbox.ru/media/code/kak-ustroeno-mashinnoe-obuchenie-zadachi-algoritmy-i-vidy-machine-learning/>. (accessed: 01.02.25) [in Russian]
2. Montgomery D.S. Introduction to Linear Regression Analysis / D.S. Montgomery, E.A. Pek, G.G. Wining. — Moscow: John Wiley & Sons Limited, 2021. — 672 p.
3. Bondarev A.V. Primenenie algoritmov mashinnogo obucheniya v prognoze rezultata piroliticheskogo analiza [Application of machine learning algorithms in result forecast pyrolytic analysis] / A.V. Bondarev, Ye.V. Kozlova // PoliPRESS. — 2021. — №3. — P. 333–336. — URL: https://www.elibrary.ru/item.asp?utm_source=chatgpt.com&id=47167532 (accessed: 02.05.25). [in Russian]
4. Sultanov B.R. Ispol'zovanie metoda ekstremal'nogo gradientnogo bustinga v prognozirovaniyu konsolidirovannogo byudzheta regiona [Using the Method of Extreme Gradient Boosting in Forecasting the Consolidated Budget of the Region] / B.R. Sultanov, D.R. Sultangareev // Economics and Management: Scientific and Practical Journal. — 2021. — № 5 (161). — P. 134–139. — URL: https://ekam-journal.com/images/6-2021/5-2021/Sultanov_Sultangareev.pdf?utm_source=chatgpt.com (accessed: 02.02.25). — DOI: 10.34773/EU.2021.5.23 [in Russian]
5. Zakharenko Ye. Gradientnii busting. Realizatsiya s nulya na Python i razbor osobennostei yego modifikatsii (XGBoost, CatBoost, LightGBM) [Gradient Boosting: Implementation from Scratch in Python and Analysis of Its Variants (XGBoost, CatBoost, LightGBM)] [Gradient Boosting: Implementation from Scratch in Python and Analysis of Its Variants (XGBoost, CatBoost, LightGBM)]

CatBoost, LightGBM)] / Ye. Zakharenko // Habr. — 2024. — URL: <https://habr.com/ru/articles/799725/> (accessed: 02.02.25) [in Russian]

6. Badanina N.V. Klasterizatsiya voditelei po stepeni opasnosti vozhdeniya s ispolzovaniem algoritmov mashinnogo obucheniya [Clustering of Drivers by Driving Risk Using Machine Learning Algorithms] / N.V. Badanina, V.A. Sudakov // Modelirovanie i analiz dannikh [Modelling and Data Analysis]. — 2022. — Vol. 12. — № 1. — P. 5–15. — URL: https://psyjournals.ru/journals/mda/archive/2022_n1/Badanina_Sudakov (accessed: 01.02.25). — DOI: 10.17759/mda.2022120101 [in Russian]

7. Ivanova A.A. Neirosetevie tekhnologii v sisteme otsenki resursnogo potentsiala agrarnogo sektora APK [Neural Network Technologies in the System for Assessing the Resource Potential of the Agricultural Sector] / A.A. Ivanova, O.Yu. Smislova // Ekonomika, predprinimatelstvo i pravo [Economics, Entrepreneurship and Law]. — 2024. — Vol. 14. — № 10. — P. 5771–5786. — URL: <https://1economic.ru/lib/121808> (accessed: 02.02.25). — DOI: 10.18334/epp.14.10.121808 [in Russian]

8. Bozhigitov B.E. Prognozirovaniye vremennyx ryadov s ispol'zovaniem modeli ARIMA [Forecasting time series using the ARIMA model] / B.E. Bozhigitov, Zh.Zh. Kozhamkulova // Foundations and Trends in Research. — 2024. — № 8. [in Russian]

9. Alzheev A.V. Sravnitelniy analiz prognoznikh modelei ARIMA i LSTM na primere aktsii rossiiskikh kompanii [Comparative Analysis of ARIMA and LSTM Predictive Models: Evidence from Russian Stocks] / A.V. Alzheev, R.A. Kochkarov // Finansy: teoriya i praktika [Finance: Theory and Practice]. — 2020. — Vol. 24. — № 1. — P. 14–23. — URL: <https://financetp.fa.ru/jour/article/view/950> (accessed: 03.02.25). — DOI: 10.26794/2587-5671-2020-24-1-14-23 [in Russian]

10. Sedukhin O. Interpretatsiya modelei i diagnostika sdvigov dannikh: LIME, SHAP i Shapley Flow [Model Interpretation and Data Shift Diagnostics: LIME, SHAP, and Shapley Flow] / O. Sedukhin // Habr. — 2022. — URL: <https://habr.com/ru/companies/ods/articles/599573/> (accessed: 31.01.25) [in Russian]

11. Matorin S.I. Informatsionno-analiticheskaya tekhnologiya budushchego – Internet veshchei (IoT) [Information-Analytical Technology of the Future – Internet of Things (IoT)] / S.I. Matorin, S.V. Gul // Nauchnyi rezultat. Informatsionnie tekhnologii [Scientific Result. Information Technologies]. — 2023. — Vol. 8. — № 4. — P. 70–78. — URL: <https://rrinformation.ru/journal/article/view/2518-1414-2023-8-4-0-6> (accessed: 03.02.25). — DOI: 10.18413/2518-1414-2023-8-4-0-6 [in Russian]

12. Chto vliyayet na stoimost' avtomobilya [What Affects the Cost of a Car] // AutoCode. — 2021. — URL: <https://avtocod.ru/chto-vliyayet-na-stoimost-avtomobilya>. (accessed: 05.02.25) [in Russian]