

# АЛГОРИТМЫ ПРОГНОЗИРОВАНИЯ СТОИМОСТИ АВТОМОБИЛЕЙ С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ

Щерба А. П., Нестеренков С. Н.

Кафедра программного обеспечения информационных технологий,  
Белорусский государственный университет информатики и радиоэлектроники  
Объединённый институт проблем информатики Национальной академии наук Беларуси  
Минск, Республика Беларусь  
E-mail: scherba891@gmail.com, s.nesterenkov@bsuir.by

*В работе рассматриваются алгоритмы прогнозирования стоимости автомобилей с использованием методов машинного обучения, этапы их развития и ключевые факторы успеха. Обсуждаются основные концепции предсказательных моделей, важность выбора признаков и обработки данных, а также использование различных алгоритмов, таких как линейная регрессия, деревья решений и нейронные сети. Проводится анализ мировых трендов и актуальных решений в области машинного обучения для оценки стоимости автомобилей в различных сегментах. Кроме того, рассматриваются инструменты разработки и программного обеспечения, поддерживающие процесс построения моделей, а также задачи создания точных и эффективных алгоритмов, способных работать с большими массивами данных в открытой экосистеме прогнозирования.*

## ВВЕДЕНИЕ

Прогнозирование стоимости автомобилей является важной задачей в автомобильной индустрии, особенно для продажи, покупки и страхования транспортных средств. Методы машинного обучения (ML) позволяют автоматизировать процесс оценки стоимости автомобилей, обеспечивая более точные и оперативные предсказания на основе множества факторов. Современные ML-алгоритмы могут анализировать большие объемы данных и выявлять скрытые закономерности, что делает их идеальными инструментами для решения этой задачи.

### I. ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ

Алгоритмы машинного обучения, такие как регрессия, деревья решений и нейронные сети, используются для анализа различных факторов, влияющих на стоимость автомобиля, таких как возраст, пробег, марка, модель, техническое состояние и другие параметры. Каждый из этих методов имеет свои преимущества и недостатки, и выбор конкретного подхода зависит от доступных данных и требуемой точности прогнозов.

- **Регрессионные методы.** Линейная регрессия и её улучшенные версии, такие как полиномиальная регрессия, позволяют строить зависимость между признаками и целевой переменной – стоимостью автомобиля. Эти методы хорошо подходят для моделирования относительно простых данных с линейными зависимостями, но могут быть недостаточны для сложных сценариев.
- **Деревья решений и ансамблевые методы.** Случайный лес и градиентный бустинг – это ансамблевые методы, которые строят предсказания на основе объединения решений нескольких деревьев решений. Эти

методы являются более гибкими и часто показывают высокую точность при работе с данными о стоимости автомобилей.

- **Нейронные сети.** Глубокие нейронные сети применяются для анализа сложных нелинейных зависимостей и могут обеспечить высокую точность прогнозов. Однако они требуют больших объемов данных и вычислительных ресурсов для обучения.

### II. ТЕХНОЛОГИЧЕСКИЕ ВОЗМОЖНОСТИ И ПРЕИМУЩЕСТВА

Применение машинного обучения в прогнозировании стоимости автомобилей предоставляет ряд уникальных преимуществ:

- **Точность предсказаний.** Методы машинного обучения позволяют учитывать множество факторов одновременно, что повышает точность прогнозов. В отличие от традиционных статистических методов, ML-алгоритмы могут выявлять сложные зависимости, которые не очевидны для человека. Кроме того, использование алгоритмов машинного обучения позволяет значительно сократить влияние субъективных факторов на процесс оценки стоимости.
- **Адаптивность.** Системы на основе машинного обучения могут адаптироваться к новым данным. Это особенно важно в условиях меняющегося рынка автомобилей, где стоимость транспортных средств может изменяться в зависимости от сезонных факторов, экономической ситуации и состояния рынка.
- **Автоматизация процесса.** Автоматизация оценки стоимости с помощью ML-алгоритмов позволяет сократить время на проведение анализа, минимизировать чело-

веческие ошибки и стандартизировать процесс оценки.

### III. ВЫЗОВЫ И ОГРАНИЧЕНИЯ

Несмотря на множество преимуществ, использование методов машинного обучения для прогнозирования стоимости автомобилей связано с рядом технологических вызовов:

- **Качество данных.** Одной из ключевых проблем является наличие и качество данных. Недостаток информации, пробелы в данных или наличие некорректных значений могут негативно повлиять на качество прогнозов. Для решения этой проблемы необходима тщательная предварительная обработка данных, включая их очистку и нормализацию.
- **Сложность моделей.** Более сложные модели, такие как нейронные сети, требуют значительных вычислительных ресурсов для обучения. Это может стать препятствием для их использования в условиях ограниченных ресурсов или при необходимости быстрого получения результатов.
- **Переобучение.** При работе с небольшими или избыточными наборами данных модели могут переобучаться, что снижает их способность к обобщению на новых данных. Для борьбы с этим применяются техники регуляризации и использование ансамблей моделей.

### IV. ПРИМЕРЫ РЕАЛЬНОГО ПРИМЕНЕНИЯ

Прогнозирование стоимости автомобилей на основе методов машинного обучения уже активно используется в различных областях, демонстрируя широкий спектр возможностей для автоматизации и повышения точности оценок.

- **Автодилеры** применяют ML-алгоритмы для автоматической оценки автомобилей на основании их технических характеристик, возраста, пробега и рыночной ситуации. Благодаря машинному обучению, процесс оценки значительно ускоряется, становится более объективным и минимизирует человеческий фактор, что особенно полезно при работе с большими объемами автомобилей.
- **Страховые компании** используют прогнозные модели для более точного определения страховых премий на основе таких факторов, как возраст автомобиля, его состояние, пробег, а также статистических данных о частоте аварий и технических неисправностях для аналогичных моделей. Это позволяет страховым компаниям точнее рассчитывать риски и предлагать индивидуализированные страховые полисы.
- **Онлайн-платформы** для продажи автомобилей (например, CarGurus, AutoTrader,

Drom) внедряют машинное обучение для предсказания рыночной стоимости автомобилей. Такие платформы анализируют историю продаж, рыночные тенденции, региональные особенности и параметры автомобилей, чтобы помочь продавцам установить справедливую цену, а покупателям — получить лучшее предложение.

### V. ЗАКЛЮЧЕНИЕ

Алгоритмы машинного обучения предоставляют мощные инструменты для прогнозирования стоимости автомобилей, делая процесс оценки более точным, быстрым и адаптивным. Современные подходы, такие как линейная регрессия, деревья решений и нейронные сети, способны учитывать многочисленные факторы, включая марку, модель, возраст автомобиля, пробег и рыночные тенденции. Эти методы помогают выявлять сложные зависимости и паттерны, которые могут быть упущены традиционными статистическими методами.

Тем не менее, несмотря на очевидные преимущества, точность алгоритмов сильно зависит от качества данных. Данные должны быть тщательно очищены и подготовлены для того, чтобы алгоритмы машинного обучения могли достичь оптимальных результатов. Более того, такие методы требуют значительных вычислительных ресурсов, особенно при работе с большими наборами данных.

С развитием технологий и увеличением объемов доступных данных, возможности машинного обучения в оценке стоимости автомобилей будут только расти. Эти алгоритмы станут незаменимыми инструментами для автодилеров, страховых компаний и частных покупателей, помогая им принимать более обоснованные и точные решения на рынке автомобилей.

### VI. СПИСОК ЛИТЕРАТУРЫ

1. Кучеренко, В. Т. Прогнозирование дневного количества осадков методами машинного обучения / В. Т. Кучеренко, С. Н. Нестеренков, И. В. Шилов, А. Н. Марков // *BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня: сб. научных статей VIII Междунар. науч.-практ. конф. (Республика Беларусь, Минск, 11-12 мая 2022 года)* / редкол.: В. А. Богущ [и др.]. – Минск : Бестпринт, 2022. – С. 320-324.
2. Нестеренков, С. Н., & Белов, К. П. (2017). Модифицированный генетический алгоритм для обучения нейронной сети. В *Информационные технологии и системы 2017 (ИТС 2017): материалы международной научной конференции, Минск, 25 октября 2017 г.*, Белорусский государственный университет информатики и радиоэлектроники, редкол.: Л. Ю. Шилин [и др.] (с. 204-205). Минск.
3. Balcioglu, Y. S., & Sezen, B. (2024). Car Price Prediction Using Machine Learning Techniques. In *Proceedings of the 6th International Artemis Congress on Health and Sport Sciences*.
4. Gron, A. (2020). The Price of Cars in the Automobile Industry: The Effect of Market Structure on Car Prices. *The RAND Journal of Economics*, 31(1), 124-139.