Министерство науки и высшего образования Российской Федерации

федеральное государственное автономное образовательное учреждение высшего образования

«Национальный исследовательский университет ИТМО»

Факультет инфокоммуникационных технологий

**ПРАКТИЧЕСКАЯ РАБОТА №6**

**Выполнил**:

студент 4 курса ИМРиП

Группы D34101

Ф.И.О. Лю Ифэн

Лю Юе

**Руководитель**:

Чернышева Анастасия Вадимовна

Санкт-Петербург

2024

**1. Формулировка задачи машинного обучения**

Задача машинного обучения, которую нам нужно решить, заключается в предсказании цены продажи подержанных автомобилей на основе их характеристик, таких как марка, модель, пробег, год выпуска и другие. Эта задача является задачей регрессии.

**2. Ограничения и требования предметной области**

Точность предсказаний: Цена должна предсказываться с высокой точностью, чтобы быть полезной для потенциальных покупателей и продавцов.

Обобщающая способность: Модель должна хорошо работать не только на данных, на которых она была обучена, но и на новых, ранее невиданных данных.

Время обучения и предсказания: Время, необходимое для обучения модели и делания предсказаний, должно быть разумным.

**3. Метрики оценки качества**

MSE (Средняя квадратичная ошибка): Используется для оценки среднего размера ошибок в предсказаниях. Чем ниже значение, тем лучше.

R² (Коэффициент детерминации): Отражает долю дисперсии зависимой переменной, объяснённую моделью. Значение близкое к 1 указывает на высокую точность модели.

Эти метрики подходят, так как они широко используются для оценки качества моделей регрессии, позволяя оценить как абсолютные ошибки предсказаний (MSE), так и относительную точность модели (R²).

**4. Оценка результатов**

По результатам ПР4: Лучший результат показал алгоритм Random Forest с MSE=4.75 и R²=0.94.

По результатам ПР5: Глубокая нейронная сеть достигла MSE=2.25 и R²=0.968.

Оба подхода показали высокую точность предсказаний, выполнив основные требования к результату.

**5. Примеры неудачных предсказаний**

В ходе прогона не было обнаружено конкретных примеров сбоя модели.

**6. Выводы и улучшение моделей**

Глубокая нейронная сеть из ПР5 показала лучшую производительность по сравнению с моделями из ПР4, предлагая наилучшее сочетание MSE и R².

Улучшение параметров: Для дальнейшего улучшения модели стоит рассмотреть оптимизацию архитектуры сети, настройку гиперпараметров (скорость обучения, размеры пакетов) и использование техник предотвращения переобучения (Dropout, регуляризация).

**Заключение**

глубокое изучение показало выдающиеся результаты в задаче прогнозирования цен на подержанные автомобили. Однако для достижения оптимальной производительности и обеспечения высокой обобщающей способности необходимо продолжать эксперименты с архитектурой модели, обработкой данных и методами обучения.

https://github.com/l769030551/used\_care\_machine\_learning