**Описание проекта**

Поздравляем! Вы прошли курс в тренажёре. Самое время проверить знания и решить новую задачу машинного обучения. Выполнять работу будете самостоятельно.

Когда закончите, отправьте её на проверку наставнику: он пришлёт комментарии в течение суток. После этого нужно доработать проект и пройти повторную проверку.

Скорее всего, вы будете дорабатывать кейс по комментариям ещё несколько раз. Это нормально.

Проект завершён, когда наставник одобрит все доработки.

**Описание проекта**

Сервис по продаже автомобилей с пробегом «Не бит, не крашен» разрабатывает приложение для привлечения новых клиентов. В нём можно быстро узнать рыночную стоимость своего автомобиля. В вашем распоряжении исторические данные: технические характеристики, комплектации и цены автомобилей. Вам нужно построить модель для определения стоимости.

Заказчику важны:

* качество предсказания;
* скорость предсказания;
* время обучения.

**Инструкция по выполнению проекта**

1. Загрузите и подготовьте данные.
2. Обучите разные модели с различными гиперпараметрами.
3. Проанализируйте скорость работы и качество моделей.

Примечания:

* Для оценки качества моделей применяйте метрику *RMSE*.
* Самостоятельно освойте библиотеку *LightGBM* и её средствами постройте модели градиентного бустинга.
* Время выполнения ячейки кода *Jupyter Notebook* можно получить специальной командой. Найдите её.
* Поскольку модель градиентного бустинга может обучаться долго, измените у неё только два-три параметра.
* Если перестанет работать *Jupyter Notebook*, удалите лишние переменные оператором del:

Скопировать кодPYTHON

del features\_train

**Описание данных**

Данные находятся в файле /datasets/autos.csv. [Скачать датасет](https://code.s3.yandex.net/datasets/autos.csv).

**Признаки**

* *DateCrawled* — дата скачивания анкеты из базы
* *VehicleType* — тип автомобильного кузова
* *RegistrationYear* — год регистрации автомобиля
* *Gearbox* — тип коробки передач
* *Power* — мощность (л. с.)
* *Model* — модель автомобиля
* *Kilometer* — пробег (км)
* *RegistrationMonth* — месяц регистрации автомобиля
* *FuelType* — тип топлива
* *Brand* — марка автомобиля
* *NotRepaired* — была машина в ремонте или нет
* *DateCreated* — дата создания анкеты
* *NumberOfPictures* — количество фотографий автомобиля
* *PostalCode* — почтовый индекс владельца анкеты (пользователя)
* *LastSeen* — дата последней активности пользователя

**Целевой признак**

*Price* — цена (евро)

**Как будут проверять проект?**

Мы подготовили критерии оценки проекта, которыми руководствуются наставники. Прежде чем приступить к решению кейса, внимательно их изучите.

На что обращают внимание наставники, проверяя проект:

* Все ли шаги по инструкции выполняете?
* Как готовите данные?
* Какие модели и гиперпараметры рассматриваете?
* Не дублируете ли код?
* Какие выводы делаете?
* Следите ли за структурой проекта?
* Поддерживаете аккуратность кода?

Всё, что вам нужно знать, есть в шпаргалках и конспектах прошлых тем.

Успеха!