Project Description

Congratulations! You've completed another training platform course. Now is the perfect time to test your skills and solve a new machine learning problem. For this project, you will be on your own.

When you finish, send your work to the project reviewer. You will receive feedback within 24 hours. After that, you will make any necessary changes to your work and send it for a second review.

Usually, this process has to be repeated several times until you get the green light from the reviewer and all the corrections are approved. That's all part of the job.

Your project will be considered complete once the project reviewer approves it.

Project description

Rusty Bargain used car sales service is developing an app to attract new customers. In that app, you can quickly find out the market value of your car. You have access to historical data: technical specifications, trim versions, and prices. You need to build the model to determine the value.

Rusty Bargain is interested in:

- the quality of the prediction
- the speed of the prediction
- the time required for training

Project instructions

1. Download and look into the data.
2. Train different models with various hyperparameters.
3. Analyze the speed and quality of models.

Notes:

- Use the *RMSE* metric to evaluate the models.

- On your own, work with the LightGBM library and use its tools to build gradient boosting models.
- You can use a special command to find the cell code runtime in Jupyter Notebook. Find that command.
- Since the training of a gradient boosting model can take a long time, change only a few model parameters.
- If Jupyter Notebook stops working, delete the excessive variables using the `del` operator:

```
del features_train
```

Data description

The dataset is stored in file `/datasets/car_data.csv`. download dataset{target="_blank"}.

Features

- *DateCrawled* — date profile was downloaded from the database
- *VehicleType* — vehicle body type
- *RegistrationYear* — vehicle registration year
- *Gearbox* — gearbox type
- *Power* — power (hp)
- *Model* — vehicle model
- Mileage — mileage (measured in km due to dataset's regional specifics)
- *RegistrationMonth* — vehicle registration month
- *FuelType* — fuel type
- *Brand* — vehicle brand
- *NotRepaired* — vehicle repaired or not
- *DateCreated* — date of profile creation
- *NumberOfPictures* — number of vehicle pictures
- *PostalCode* — postal code of profile owner (user)

- *LastSeen* — date of the last activity of the user

Target

Price — price (Euro)

Project evaluation

We've put together the evaluation criteria for the project. Read this carefully before moving on to the task.

Here's what the reviewers will look at when reviewing your project:

- Have you followed all the steps of the instructions?
- How did you prepare the data?
- What models and hyperparameters have you considered?
- Have you managed to avoid code duplication?
- What are your findings?
- Have you kept to the project structure?
- Have you kept the code neat?

You have your takeaway sheets and chapter summaries, so you are ready to proceed to the project.

Good luck!