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

Beta Bank customers are leaving: little by little, chipping away every month. The bankers figured out it's cheaper to save the existing customers rather than to attract new ones.

We need to predict whether a customer will leave the bank soon. You have the data on clients' past behavior and termination of contracts with the bank.

Build a model with the maximum possible *F1* score. To pass the project, you need an *F1* score of at least 0.59. Check the *F1* for the test set.

Additionally, measure the *AUC-ROC* metric and compare it with the *F1*.

Project instructions

1. Download and prepare the data. Explain the procedure.

2. Examine the balance of classes. Train the model without taking into account the imbalance. Briefly describe your findings.

3. Improve the quality of the model. Make sure you use at least two approaches to fixing class imbalance. Use the training set to pick the best parameters. Train different models on training and validation sets. Find the best one. Briefly describe your findings.

4. Perform the final testing.

Data description

The data can be found in `/datasets/Churn.csv` file. <u>Download the dataset.</u>

Features

- *RowNumber* — data string index
- *CustomerId* — unique customer identifier
- *Surname* — surname
- *CreditScore* — credit score
- *Geography* — country of residence
- *Gender* — gender
- *Age* — age
- *Tenure* — period of maturation for a customer's fixed deposit (years)
- *Balance* — account balance
- *NumOfProducts* — number of banking products used by the customer
- *HasCrCard* — customer has a credit card
- *IsActiveMember* — customer's activeness
- *EstimatedSalary* — estimated salary

Target

- *Exited* — customer has left

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:

- How did you prepare the data for training? Have you processed all of the feature types?

- Have you explained the preprocessing steps well enough?

- How did you investigate the balance of classes?

- Did you study the model without taking into account the imbalance of classes?

- What are your findings about the task research?

- Have you correctly split the data into sets?

- How have you worked with the imbalance of classes?

- Did you use at least two techniques for imbalance fixing?

- Have you performed the model training, validation, and final testing correctly?

- How high is your *F1* score?

- Did you examine the *AUC-ROC* values?

- Have you kept to the project structure and kept the code neat?

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

Good luck!

Let's go!

Or do the project on your computer and upload your work when you're done. **Upload and submit** ↑

Next