

Задание № 4. Kaggle соревнование.

Нужно решить одно из предложенных соревнований:

Предсказание клиентов, склонных к оттоку (20) – соревнование для новичков. Это игрушечное соревнование для тех, кто никогда не решал Kaggle. Используется известный датасет Telco оттока пользователей провайдера. Можно объединяться в команды из **2ух** человек. Это соревнование дает **20 баллов**.

<https://www.kaggle.com/t/3eb01fc2677b49b7a8c62c059aa749a5>

Предсказание клиентов, которым не стоит выдавать кредит – активное соревнование для уверенных в своих силах. Можно объединяться в команды из **4ех** человек. Это соревнование дает **40 баллов**.

<https://www.kaggle.com/competitions/home-credit-default-risk/overview>

Дедлайн задания – **18 декабря**

Требования к сдаче

- Ссылка на решение на лидерборде.
- Оформленный ноутбук финального решения. В ноутбуке должен быть продемонстрирована метрика на **public** сете.
- Презентация. Презентация на 5-10 минут с вопросами. В презентации должны быть следующие слайды:
 - Кто чем занимался в соревновании.
 - Какую задачу решаем? На что будем обращать внимание в первую очередь при решении? Какие есть нюансы есть в задаче? Предварительный статистический анализ пользователей тренировочной выборки:
 - **Предсказание оттока** – какими отличительными характеристиками обладают клиенты склонные к оттоку?
 - **Выдача кредита** – какие отличительные характеристики есть у клиентов не склонных отдавать кредит?
 - Этапы получения финального решения:
 - Анализ данных
 - Предобработка
 - Описание финальной модели
 - Как вы будете использовать модель при запуске прод? Можно ли сделать первоначальную оценку эффективности внедрения вашей модели? За какими метриками вы будете следить?