

Сроки: 3 дня (должно занять 3-5 часов)

Цель: Задачей этого челленджа является оценка Ваших навыков анализа данных, создания и валидации модели, а также способности выбрать правильное решение.

Объем:

1. Понимание данных: Ознакомьтесь с датасетом и попробуйте проинтерпретировать каждый столбец.

2. Генерация фичей: Какие новые фичи Вы рассчитали бы из существующих, чтобы лучше отразить основные закономерности в данных?

3. Важность фичей: Выделите основные признаки в данных. Почему эти признаки важны? Как они влияют на целевую переменную?

4. Построение и оценка модели: Используя данные, создайте прогнозную модель, предсказывающую целевую переменную. Какую модель Вы выбрали и почему? Насколько хорошо работает Ваша модель?

5. Презентация: Четко и понятно опишите процесс Вашего решения и результаты работы. Можете ли Вы объяснить сложные концепции из статистики и Data Science простыми словами?

Ожидаемый результат: В качестве решения необходимо прислать ссылку на Github с выкладками в Jupyter Notebook.