

## Задание 4. Анализ данных и обработка признаков Base

Курс по методам машинного обучения, 2025-2026, Богачев Владимир

### 1 Характеристики задания

Уровень Base

- **Длительность:** 1 неделя
- **Кросс-проверка:** 10 баллов; в течение 1 недели после дедлайна; Нельзя сдавать после дедлайна.

Почта

- **Почта:** ml.cmc@mail.ru
- **Темы для писем на почту:** ВМК.ML[Задание 4][peer-review]

**Кросс-проверка:** После окончания срока сдачи, у вас будет еще неделя на проверку решений как минимум **3х других студентов** — это **необходимое** условие для получения оценки за вашу работу. Если вы считаете, что вас оценили неправильно или есть какие-то вопросы, можете писать на почту с соответствующей темой письма.

### 2 Описание задания

В задании вы познакомитесь с методами анализа и обработки данных, а также узнаете, каким способом можно использовать категориальные признаки в машинном обучении.

### 3 Кросс-проверка

**Внимание!** Отправлять задание нужно в систему во вкладку с пометкой (notebook).

**Внимание!** Отправлять задание нужно только с расширением `ipynb`! После отправки проверьте корректность загруженного задания в систему, просмотрев глазами загруженное решение (оно автоматически конвертируется в `html`). Как это сделать, можно найти в tutoriale по проверяющей системе на сайте курса.

**Внимание!:** Перед сдачей проверьте, пожалуйста, что не оставили в ноутбуке где-либо свои ФИО, группу и так далее — кросс-рецензирование проводится анонимно.