

# Домашнее задание №2 по курсу «Методы машинного обучения»

Терлиженко Ирина

26 февраля 2020 г.

Необходимо реализовать и исследовать на практике несколько feature selection алгоритмов. В качестве датасета использовать Boston dataset . В качестве предсказываемого признака (метки) используется признак MEDV.

В задании 1 все необходимые вычисления (кроме скачивания датасета) необходимо делать без использования каких-либо ML библиотек (sklearn, pandas можно использовать только для визуализации). В остальных задачах разрешено использование библиотек кроме методов, реализующих непосредственно feature selection.

На проверку необходимо прислать feature\_selection.ipynb файл с кодом и отчетом и/или feature\_selection.py файл с кодом и .pdf файл с отчетом.

## 1 Задания

1. (+2) Выберите 5 лучших признаков на основании Filter метода. В качестве меры качества используйте коэффициент корреляции Пирсона.
2. Выберите 5 лучших признаков на основании Forward (+1) и Backward (+1) Greedy Selection методов. В качестве базового алгоритма используйте линейную регрессию с квадратичной функцией потерь.
3. (+2) Выберите 5 лучших признаков, используя линейную регрессию с квадратичной функцией потерь и регуляризацией Лассо. В отчете опишите, как вы выбирали параметр регуляризации  $\lambda$ .