

Машинное обучение

ФКН ВШЭ

Теоретическое домашнее задание №2

Задача 1. Найдите константу C , решающую следующую задачу ($0 < \tau < 1$ фиксировано):

$$\sum_{i=1}^{\ell} \rho_{\tau}(y_i - C) \rightarrow \min_C,$$
$$\rho_{\tau}(x) = \begin{cases} \tau x, & x > 0, \\ (\tau - 1)x, & x \leq 0. \end{cases}$$

Задача 2. Пусть функция потерь задается следующим образом ($\varepsilon > 0$ фиксировано):

$$L(y_i, a(x_i)) = \varphi_{\varepsilon}(y_i - x_i^T w),$$
$$\varphi_{\varepsilon}(t) = \begin{cases} 0, & -\varepsilon < t < \varepsilon, \\ t - \varepsilon, & t > \varepsilon, \\ -t - \varepsilon, & t < -\varepsilon. \end{cases}$$

Сведите задачу оптимизации

$$\sum_{i=1}^{\ell} \varphi_{\varepsilon}(x_i^T w - y_i) \rightarrow \min_w$$

к задаче линейного программирования.

Задача 3. Убедитесь, что вы знаете ответы на следующие вопросы:

- Почему L_1 -регуляризация производит отбор признаков?
- Почему коэффициент регуляризации нельзя подбирать по обучающей выборке?
- Что такое кросс-валидация, чем она лучше использования отложенной выборки?
- Почему категориальные признаки нельзя закодировать натуральными числами? Что такое one-hot encoding?

-
- Для чего нужно масштабировать матрицу объекты-признаки перед обучением моделей машинного обучения?
 - Почему MSE чувствительно к выбросам?
 - Какие методы можно применять для оптимизации MAE?
 - Что такое Huber Loss? В чем его преимущества по сравнению с MAE и MSE?
 - Почему квантильная регрессия так называется?
 - Какова общая постановка задачи линейного программирования?