

Машинное обучение.

Домашнее задание №4

Задача 1. Пусть базовые алгоритмы из семейства \mathcal{A} возвращают ответы из множества $\mathbb{Y} = \{-1, +1\}$. Покажите, что в случае с экспоненциальной функцией потерь

$$Q(a) = \sum_{i=1}^{\ell} \exp(-y_i a(x_i))$$

алгоритм AnyBoost в точности совпадает с AdaBoost.

Задача 2. Пусть известно распределение на объектах и ответах $p(x, y)$. Ответы на объектах принадлежат множеству $\mathbb{Y} = \{-1, +1\}$. Рассмотрим алгоритм, который возвращает ответ, минимизирующий матожидание экспоненциальной функции потерь в данной точке:

$$a^*(x) = \arg \min_{a \in \mathbb{R}} \mathbb{E}_{y|x} e^{-ya} = \arg \min_{a \in \mathbb{R}} \sum_{y \in \{-1, +1\}} p(y|x) e^{-ya}.$$

Найдите в явном виде алгоритм $a^*(x)$.

Задача 3. Выпишите формулы для поиска базового алгоритма b_N и коэффициента γ_N в градиентном бустинге с экспоненциальной функцией потерь.