

Правила

Работа выполняется самостоятельно. Вы можете пользоваться любыми источниками, но не можете общаться с другими людьми в процессе написания работы. Пожалуйста, пишите подробные решения!

Желаем удачи!

Задача 1. (7+7 баллов) Рассмотрим выборку $(3, 0), (6, 0), (9, 0), (3, 3), (6, 3), (9, 3)$. Найти иерархию кластеров, которую вернёт иерархическая агломеративная кластеризация (bottom-up hierarchical clusterization) при обработке этой выборки. В качестве метрики на точках использовать

$$d((x_1, x_2), (x'_1, x'_2)) = 2|x_1 - x'_1| + |x_2 - x'_2|.$$

В качестве метрики на кластерах использовать

а. наименьшее расстояния между точками кластера:

$$d(A, B) = \min_{x \in A, y \in B} d(x, y).$$

б. наибольшее расстояние между точками кластера:

$$d(A, B) = \max_{x \in A, y \in B} d(x, y).$$

Если несколько кластеров находятся на одинаковом наименьшем расстоянии, выбрать произвольным образом, какие кластеры объединять. В ответе привести разбиение исходных точек на два самых больших кластера.

Задача 2. (20 баллов) Рассмотрим задачу регрессии, решаемую с помощью градиентного бустинга. Пусть данные устроены следующим образом:

x	y
0	0
3	14
2	6

Наш алгоритм строится как последовательность:

$$a_{i+1}(x) = a_i(x) + \eta_i b_i(x),$$

где $a_0(x)$ — линейная регрессия без свободного члена, $b_i(x)$ — решающие пни, $i = 0, \dots$. Функция потерь во всех случаях среднеквадратичная.

Пусть сделан один шаг градиентного бустинга, то есть мы рассматриваем модель

$$a_1 = a_0(x) + \eta_0 b_0(x).$$

Пусть b_0 обучается на антиградиент функции потерь и $\eta_0 = 1/2$.

Найти предсказания модели для всех значений x из обучающей выборки.

Задача 3. (20 баллов) Рассмотрим рекуррентную нейросеть, заданную следующим образом:

$$h_{i+1} = \text{ReLU}(w_h h_i + w_x x_{i+1}), \quad i = 0, 1, \dots, \quad h_0 = 0.$$

Она обрабатывает последовательность входных данных $x_1, x_2, \dots, x_i \in \mathbb{R}^1$. Размерность внутреннего состояния h_i равна 1, w_h и w_x — веса.

Пусть входная последовательность имеет длину 3 и равна $(-3, 2, 2)$. Пусть также $w_x = 2$ и $w_h = 1$. Найти $\partial h_3 / \partial w_x$.

Задача 4. (20 баллов) Рассмотрим задачу регрессии. Выборка имеет следующий вид:

x	y
-1	-3
-1	-2
1	5

Будем обучать на ней следующий ансамбль. Из исходной выборки сделаем N новых выборок с помощью бэггинга (выборка элементов с возвращением, размер новой выборки такой же, как исходной). На каждой новой выборке обучим базовый алгоритм. Предсказание ансамбля на новом наблюдении x является средним арифметическим предсказаний всех базовых алгоритмов на этом наблюдении. Пусть базовым алгоритмом является метод одного ближайшего соседа.

К чему будет стремиться предсказание ансамбля для $x = -1$ при $N \rightarrow \infty$ по вероятности?

Задача 5. (25 баллов) Рассмотрим многослойный персептрон с одним скрытым слоем. Пусть все функции активации являются тождественными ($f(x) = x$), размерность входного пространства равна 2, размерность скрытого слоя 1, размерность выходного пространства равна 2. Констант (bias) нет. Обозначим результат работы персептрона на входном векторе $x \in \mathbb{R}^2$ через $F(x)$. Пусть функция потерь для наблюдения x имеет вид

$$L(x) = \|F(x) - x\|^2,$$

то есть в качестве целевых значений используются сами значения x , метрика стандартная евклидова.

Пусть обучающая выборка имеет вид

x_1	x_2
2	3
3	2
-2	-3
-3	-2

Допустим, нам удалось подобрать веса нейросети таким образом, чтобы сумма потерь по всей выборке достигла своего минимума. Найти $F(2, 3)$.