

Фамилия:.....

Имя:.....

Группа:.....

### Задача №1

У вас имеются следующие данные:

Наблюдение $i$	1	2	3	4	5	6
$Y_i$	9	1	6	7	7	6
$X_{1i}$	0	0.125	-0.375	0	0.175	-0.125
$X_{2i}$	0.25	-0.75	0.5	-0.25	0	0
$T_i$	1	0	0	1	1	0
Разбиение выборки	I			II		

1. Для прогнозирования целевой переменной  $Y_i$  с помощью нормализованных в предыдущем пункте признаков  $X_{1i}$  и  $X_{2i}$  вы используете метод ближайших соседей с метрикой расстояния Манхэттен. Используя двухчастную кросс-валидацию (разбиение указано в таблице), ориентируясь на значение MAE, сделайте выбор между 1 и 2 соседями.
2. Считайте, что признак  $X_{2i}$  более не доступен в данных и в качестве контрольной переменной вы рассматриваете лишь  $X_{1i}$ . Используя метод ближайших соседей с оптимальным числом соседей (выбранным исходя из результатов предыдущего пункта) оцените с помощью T-learner условный средний эффект воздействия переменной воздействия  $T_i$  на целевую переменную  $Y_i$  для 2-го наблюдения в выборке.

**Подсказка:** При внутривыборочном прогнозировании методом ближайших соседей само наблюдение рассматривается в качестве ближайшего соседа самого себя.

## Задача №2

Имеется нейросеть, включающая всего одно наблюдение по одному признаку  $x = 1$ . Значение целевой переменной равняется  $y = 10$ . Имеется лишь один скрытый слой с двумя нейронами. В качестве функции активации в скрытом и выходном слоях используется ReLU. Применяется квадратичная функция потерь. В нейросети нет смещений (констант) и все ее параметры равняются 2.

1. Изобразите графически описанную нейросеть.
2. Рассчитайте значение функции потерь данной нейросети при заданных значениях весов.
3. Рассмотрим вес, с которым признак входного слоя входит в первый нейрон. Найдите значение данного веса после одной итерации алгоритма градиентного спуска со скоростью обучения  $\alpha = 0.125$ .
4. Вы добавили в скрытый слой исключение (dropout). Нейроны отключаются независимо друг от друга с вероятностью 0.5. Повторите предыдущий пункт, определяя математическое ожидание обновленного веса, с которым признак входит в первый нейрон..

## Задача №3

У вас имеются следующие данные:

Наблюдение $i$	1	2	3	4
$Y_i$	0	12	24	36
$X_{1i}$	1	1	0	1
$X_{2i}$	1	1	1	0

Дана квадратичная функция потерь  $L(Y_i, \hat{Y}_i) = (Y_i - \hat{Y}_i)^2$ . Для прогнозирования вы используете градиентный бустинг со скоростью обучения 0.5 и одной итерацией. В качестве базовой модели используется метод наименьших квадратов лишь с константой. Для прогнозирования градиентов используется регрессионное дерево глубины 1, в качестве критерия разбиения использующего средневзвешенную дисперсию (среднеквадратическая ошибка).

1. Запишите прогнозы базовой модели.
2. С помощью одной итерации градиентного бустинга спрогнозируйте значения целевой переменной для всех наблюдений в выборке.
3. Вы решили рассмотреть скорость обучения  $\alpha \in R$  обученного в предыдущем пункте градиентного бустинга в качестве гиперпараметра. Подберите оптимальное значение скорости обучения (тюнинг), ориентируясь на критерий MSE, рассчитанный на следующей валидационной выборке:

Наблюдение $i$	1	2
$Y_i$	18	58
$X_{1i}$	1	1
$X_{2i}$	1	0