

Бонусное домашнее задание.

1. Постройте регрессионное дерево для прогнозирования y с помощью x на обучающей выборке:

y_i	x_i
100	1
102	2
103	3
50	4
55	5
61	6
70	7

Критерий деления узла на два – минимизация MSE. Узлы делятся до тех пор, пока в узле остаётся больше двух наблюдений.

2. Постройте пример, когда индекс Джини больше 0.9999.
3. Машка пять дней гадала на ромашке, а затем выкладывала очередную фотку “Машка с ромашкой” в инстаграмчик. Результат гадания – переменная y_i , количество лайков у фотки – переменная x_i . Постройте классификационное дерево для прогнозирования y_i с помощью x_i на обучающей выборке:

y_i	x_i
плюнет	10
поцелует	11
поцелует	12
к сердцу прижмёт	13
к сердцу прижмёт	14

Дерево строится до идеальной классификации. Критерий деления узла на два – максимальное падение индекса Джини.

+ решить теоретические задачи из ноутбука 8 занятия.