

## Quiz 2: на природу, в лес, к деревьям!

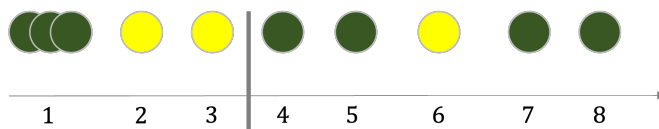
Вся наша деятельность – это борьба с энтропией. Процесс этот, понятное дело, бесконечный.

*Аркадий Волож*

Решите все задания. Все ответы должны быть обоснованы. Решения должны быть прописаны для каждого пункта. Рисунки должны быть чёткими и понятными. Все линии должны быть подписаны.

### [2] Задание 1

Посчитайте хаотичность разбиения вершины. В качестве критерия информативности используйте критерий Джини.



### [5] Задание 2

Гантер утомился видеть постоянно бездельничающих друзей на диване и решил себя развлечь, записывая, сколько чашек кофе они заказывают в неделю.  $y$  - количество чашек кофе,  $x_1$  - присоединился ли Росс к друзьям или нет,  $x_2$  - погода в градусах по Цельсию.

Гантер построил решающее дерево, прогнозирующее количество заказанных чашек кофе, и хотел было применить его, но Рейчел разлила на листок с деревом кофе. Помогите Гантеру восстановить дерево.

Дерево строится до идеального прогноза. В качестве критерия для разбиения узла Гантер использовал MSE.

$y$	$x_1$	$x_2$
12	0	5
3	0	6
3	1	7
6	1	8
100	0	13

В выборке есть выброс. Как с ним повело себя дерево? Предложите на базе этого алгоритм, который мог быть искать аномалии. Придумайте для него пафосное название.

### [3] Задание 3

Найдите дивергенцию Кульбака-Лейблера, если она определена,

- из биномиального  $\text{Bin}(\frac{1}{3}, 2)$  в равновероятное на  $0, 1, 2$ ;
- из равновероятного на  $0, 1, 2$  в биномиального  $\text{Bin}(\frac{1}{3}, 2)$ ;
- Правда ли, что она оказалась одинаковой? Почему?

### [1] Задание 4

Я сейчас читаю очень странную книгу Крис Краус "I LOVE DICK". А какую книгу читаешь ты? Учебники не считаются!