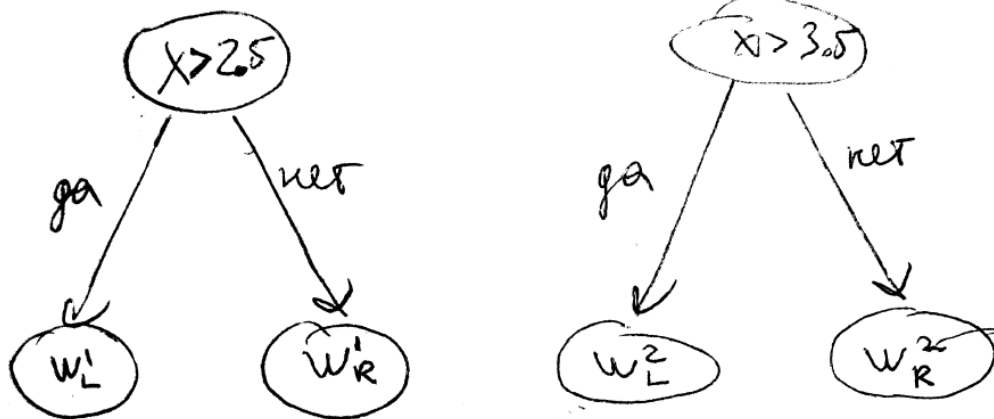


1. Задача на бэггинг

Машин-лёрнер Василий лично раздобыл выборку из четырёх наблюдений.

x_i	1	2	3	4
y_i	6	6	12	18

Два готовых дерева для леса Василий подглядел у соседа:



Василий решил использовать бэггинг. Первому дереву достались наблюдения номер 1, 1, 2 и 3. А второму дереву — 2, 3, 4 и 4. Прогнозы в каждом листе Василий строит минимизируя сумму квадратов ошибок.

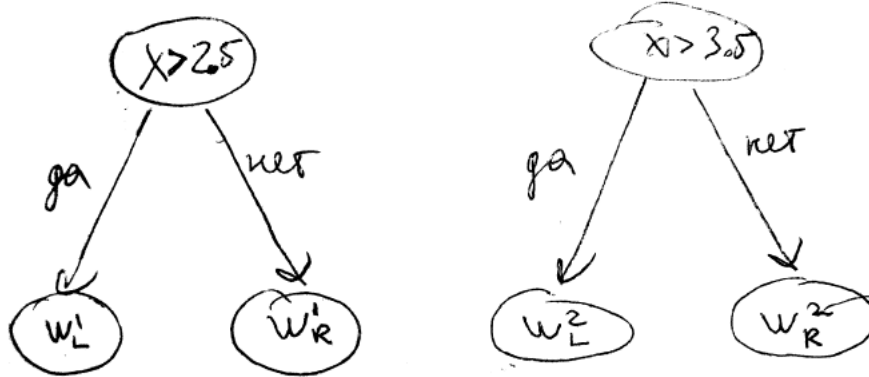
Какие прогнозы получит Василий внутри обучающей выборки с помощью своего леса?

2. Задача на бустинг

Машин-лёрнер Василий лично раздобыл выборку из четырёх наблюдений.

x_i	1	2	3	4
y_i	6	6	12	18

Два готовых дерева для бустинга Василий подглядел у соседа:



Василий решил использовать бустинг с темпом обучения η . Прогнозы в каждом листе конкретного дерева Василий строит минимизируя функцию:

$$Q = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^T w_j^2,$$

где y_i — прогнозируемое значение для i -го наблюдения, n — количество наблюдений, w_j — прогноз в j -ом листе, T — количество листов на дереве.

1. Какие прогнозы внутри обучающей выборки получит Василий при $\eta = 1$ и $\lambda = 1$?
2. Какие прогнозы внутри обучающей выборки получит Василий при $\eta = 0.5$ и $\lambda = 1$?

Домашнее задание.

1) Задача на бутстрэп

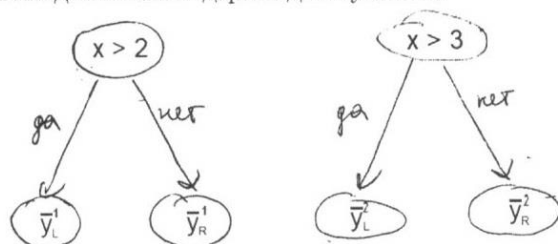
У Винни-Пуха есть 100 песенок (кричалок, вопелок, пыхтелок и сопелок). Каждый день он выбирает и поёт одну из них равномерно наугад. Одну и ту же песенку он может петь несколько раз. Сколько в среднем песенок оказываются неспетыми за 100 дней?

2) Задача на бэггинг и бустинг

Дана выборка из четырех наблюдений

x_i	1	2	3	4
y_i	6	9	9	15

Есть два готовых дерева для бустинга:



Используем Бустинг с темпом обучения η . Прогнозы в каждом листе конкретного дерева строим, минимизируя функцию:

$$Q = \sum_{i=1}^n (y_i - \bar{y}_i)^2 + \lambda \sum_{i=1}^T \bar{y}_i^2,$$

где \bar{y}_i - прогноз в i -м листе, T - количество листов на дереве.

1. Какие прогнозы внутри обучающей выборки мы получим при $\lambda = 1$ и $\eta = 1$?
2. Какие прогнозы внутри обучающей выборки мы получим при $\lambda = 1$ и $\eta = 0.5$?
3. Пусть теперь строится бэггинг, причем первое дерево строится по наблюдениям 1, 2, 4, 4, а второе – по наблюдениям 1, 2, 2, 3. Прогнозы в каждом листе строятся с помощью минимизации среднеквадратичной ошибки.

Какие прогнозы мы получим внутри обучающей выборки?