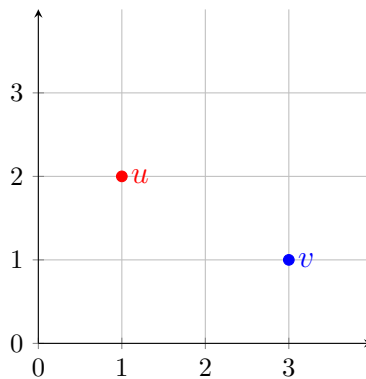


# Контрольная работа 1. Демонстрационный вариант

26 октября 2025 г.

## 1 В далекой-далекой галактике

Даны два вектора  $\mathbf{u} = (1 \ 2)$  и  $\mathbf{v} = (3 \ 1)$ . Рассчитайте манхэттенское расстояние между ними. Изобразите его на рисунке ниже.



## 2 Та-ра-рали, та-ра-ра!

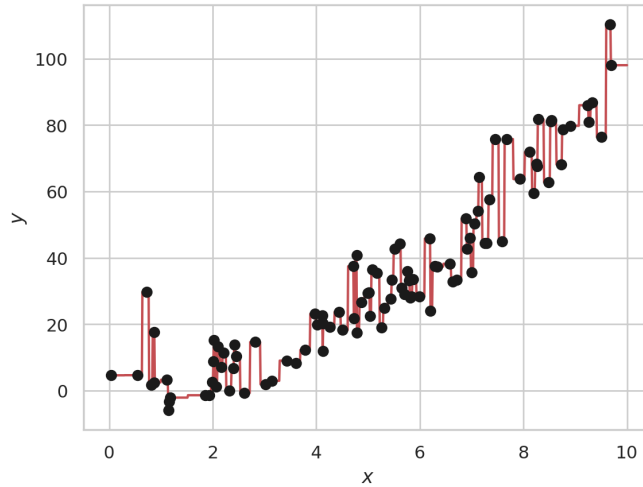
У старинушки три сына:  
Старший умный был детина,  
Средний сын и так и сяк,  
Младший вовсе был дурак.

---

Петр Ершов, Конек-горбунок

На рисунке ниже изображены обучающая выборка и модель, обученная на ней для решения задачи регрессии. Можно сказать, что данная модель

- недообучена
- нормальная
- переобучена



Какие методы борьбы с переобучением Вам известны?

### 3 Never play bongos walking down the stairs!

Рассмотрим модель линейной регрессии

$$\hat{y}_i = w_0 + w_1 x_{i1} + w_2 x_{i2} \quad (1)$$

с функцией потерь MSE (Mean Squared Error)

$$L = \text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (2)$$

Градиент функции потерь по параметрам модели имеет вид

$$\nabla_{\mathbf{w}} L = -\frac{2}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \begin{pmatrix} 1 \\ x_{i1} \\ x_{i2} \end{pmatrix}. \quad (3)$$

Сделайте один шаг градиентного спуска, если известно, что  $w_0 = w_1 = w_2 = 1$ , а скорость обучения  $\eta = 0.1$ .

$x_{i1}$	$x_{i2}$	$y_i$
3	2	1
3	3	2
4	3	3
2	4	2

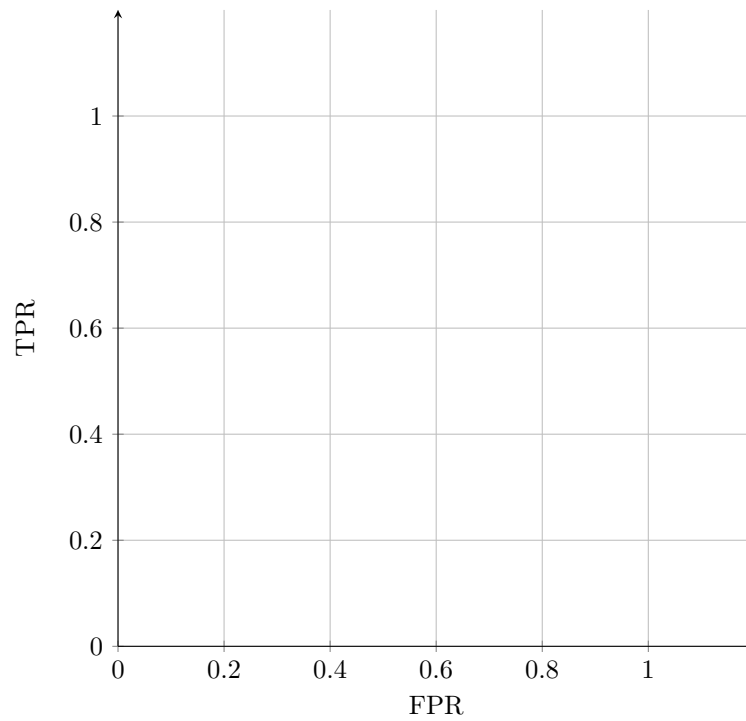
### 4 Raised on ROC

В таблице ниже представлены предсказания классификатора для набора объектов и соответствующие целевые переменные. Постройте ROC-кривую и рассчитайте AUC ROC. Как изменится AUC ROC, если из всех предсказаний вычесть 4.2?

$y_i$	1	0	0	0	1	1	0	1
$\sigma(\langle \mathbf{w}, \mathbf{x}_i \rangle)$	0.65	0.42	0.27	0.68	0.81	0.93	0.75	0.72

$$\text{TPR (True Positive Rate)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{FPR (False Positive Rate)} = \frac{\text{FP}}{\text{FP} + \text{TN}} \quad (5)$$



## 5 Ты хочешь баобаб что ли найти?!

Глаза боятся, руки делают

---

Крош, Смешарики

Какое из двух разбиений вершины дерева следует выбрать? В качестве критерия информативности (*impurity*) используйте индекс Джини.

$$H = \sum_k p_k(1 - p_k) \quad (6)$$

